WHITE PAPER REPORT:

FULL SPECTRUM TOOLS FOR COLLECTING, ANALYZING, AND USING CULTURAL DATA IN DECISION MAKING

Marvin Cohen, Rick Grannis[1], Elan Freedy, Amos Freedy and Gershon Weltman
Perceptronics Solutions, Inc., Arlington, VA

INTRODUCTION

Motivation: Problem and Approach

This paper describes the Integrated Social Network Decision Model (ISDM) as a set of support tools for collecting, analyzing, and exploiting data about cultural diverse populations. The ISDM supports full, iterative planning cycles for SSTR, disaster relief and humanitarian missions in cross cultural environments. Its purpose is to enhance data collection, data analysis, and planning aspects of such missions by a synthesis of sampling, cognitive, statistical, and decision theoretic technologies.

In today's military environment soldiers may be required to plan and conduct a broad variety of *non-kinetic* as well as conventional kinetic missions. For example, in Stability, Security, Transition, and Reconstruction (SSTR) missions, achieving the desired effects frequently involves changing both the attitudes and behaviors of a population, and success depends on leveraging their existing beliefs, values, and avenues of influence. The presence of cross-cultural human terrain adds layers of uncertainty to these tasks. Soldiers may be interacting with different sub-populations with different values, customs, and beliefs, and whose relationships with one another and with US forces are ambiguous and fluid. Ideally, soldiers would have ready access to local area experts in such matters and would themselves accumulate expertise and intuitive ability over a lengthy period of training and experience. Limits on time and resources almost always make this impossible. Soldiers who are deployed in unfamiliar surroundings and charged with delicate operations in civilian populations are nonetheless expected to "hit the ground" running. The purpose of the work described here is to provide as much computer-based support as possible for this task, with an eye to training intuition rather than replacing it.

In order to be fully effective, support tools such as the presently described ISDM must resolve a number of technical issues. These issues include enhancing the soldiers' ability to: (1) make cogent observations and collect cultural data virtually from scratch; (2) use these data to identify the relevant subpopulations involved; (3) determine the cultural elements that differentiate subpopulations from one another at different levels of resolution; (4) identify central players and map out interactions and overlapping memberships among different groups; and finally (5) use these data and analytical results for practical planning and evaluation of alternative courses of action. The present paper addresses these technical issues and outlines a set of support tools for mission planning based on minimal prior knowledge and imperfect data in cross-cultural environments.

[1] Department of Sociology, University of California Los Angeles

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **10 DEC 2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE **Full Spectrum Tools for Collecting, Analyzing, and Using Cultural Data in Decision Making**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Perceptronics Solutions, Inc,1001 19th St. N Suite 1500,Arlington,VA,22209**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**49**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

The Integrated Social Network Decision Model (ISDM) Concept

As shown in Figure 1 below, the overall ISDM concept involves two independently functioning but mutually supporting components: (1) A subsystem for the collection, analysis, and representation of demographic, social-network, and cultural data for use in decision making (right side of diagram); and (2) a subsystem for planning and decision making based on demographic, social network, and cultural data (left side of diagram). The latter in turn helps identify and evaluate further requirements for collection, analysis, and representation of data.

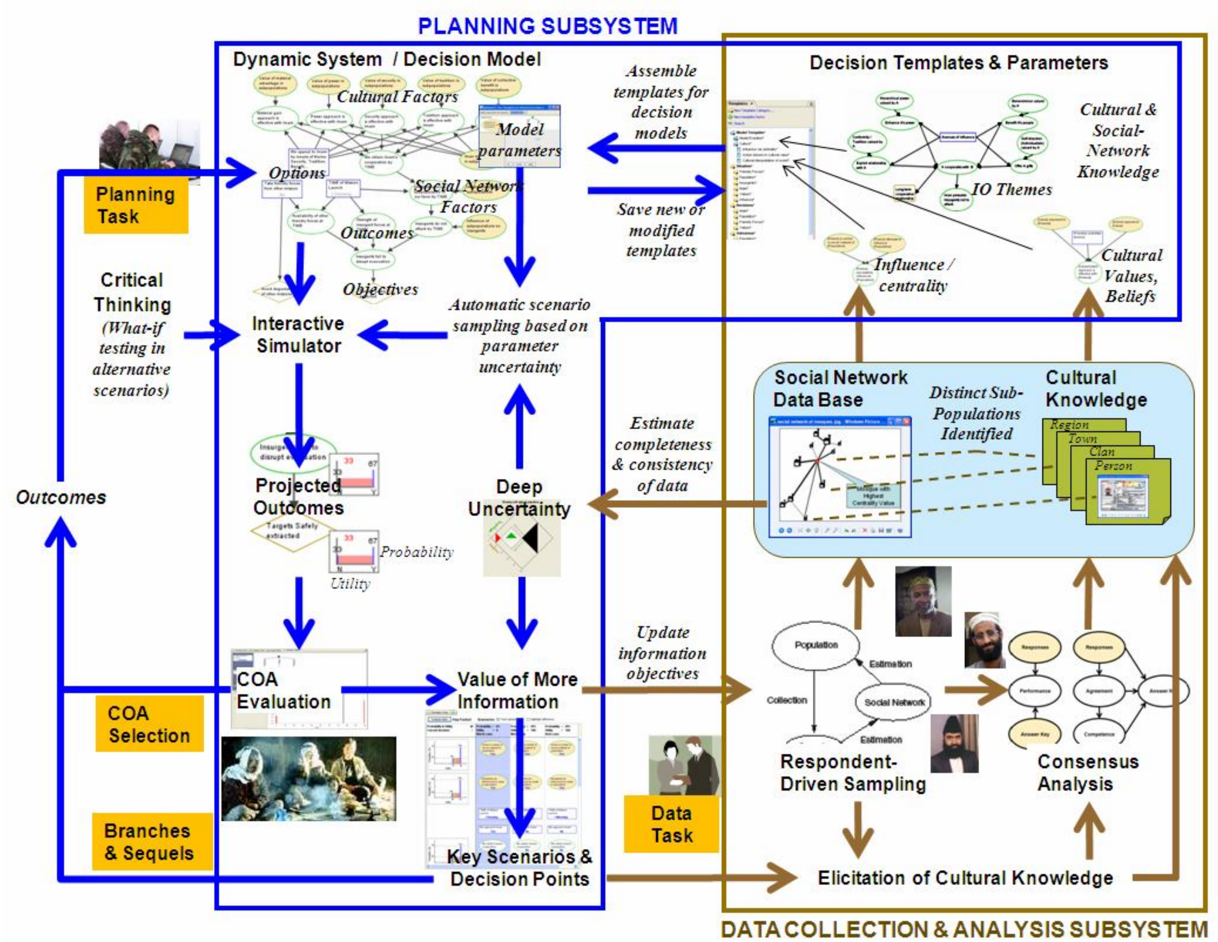

Figure 1. ISDM System Concept

The **Data Task** begins with sampling, elicitation, and analysis (lower right in Figure 1), and ends with representation (upper right). Its objectives may be influenced by a new or continuing planning process. ISDM uses a formal and empirical framework called *Respondent-Driven Sampling* (RDS) to select participants, tracks paths in the social network of target populations, and corrects results for sampling bias. A combination of semi-structured ethnographic interviews and rigorous survey methods is used to elicit respondent beliefs, values, and behaviors, including inferential relationships implicit in *mental models.* ISDM uses an independently developed formal and empirical framework called *Cultural Consensus Analysis* to extract shared cultural knowledge, identify divergent subpopulations and submodels, and assess the cultural representativeness of respondents with respect to knowledge at each level. Finally (upper right in Figure 1), ISDM uses *influence diagram*s and Bayesian decision theoretic technology for simulation and planning to represent the implications of cultural knowledge for inferences, expectations, and decisions by subject populations.. Influence diagram *templates* are organized and indexed and made available for decision modeling and iterative improvements of pre-existing models.

The **Planning Task** involves a continuous loop of situation assessment, course of action evaluation, action, and outcomes on the left side of Figure 1. Planners use a decision model (upper left) to think critically about the situation, identify potential solutions, project outcomes of alternative assumptions and courses of action, evaluate options, and anticipate new decisions down the road, including contingency plans (branches) and follow-up actions (sequels). Decisions (lower left) result in observable outcomes, which prompt further cycles of decision making. ISDM embeds this loop within a larger one, which supplies data and analysis to support planning and uses planning to direct the collection and analysis of further data. Planners assemble a decision model with the help of both general and special-purpose influence diagram templates, including representations of social network characteristics and cultural knowledge shared by specific relevant subpopulations (upper right). Concurrently, sensitivity analyses based on the current decision model estimate the value of additional information, and these measures are used to regulate further data collection and analysis (lower right), which in turn help refine or modify the plan, monitor its execution and impact, and develop follow up operations.

Our choice of the three technologies to be integrated in ISDM (Respondent Driven Sampling, Cultural Consensus Analysis, and Bayesian influence diagrams) reflects the conviction that culturally aware decision making has to break with any lingering assumption that culture is a free-floating, monolithic abstraction. The novel combination of Respondent Driven Sampling (RDS) and Cultural Consensus Analysis was suggested by team member Prof. Rick Grannis as a result of his prior work with hidden populations (Grannis, 2009; Grannis, 2010).

RDS grounds culture in individual minds connected to varying degrees by webs of social interaction with other minds; CCA identifies culturally divergent subpopulations within such networks, highlighting both shared and discrepant cultural elements. Influence diagrams show how ideas transmitted from one group to another may be tempered by assimilation into pre-existing knowledge structures, or amplified by inferential conclusions generated by interaction between old knowledge and new information. Integrated within ISDM, the three technologies provide a roadmap of cultural potentiality that can be exploited for purposes of either stability or change by planners of Information Operations, SSTR missions, etc.

RESPONDENT DRIVEN SAMPLING FROM HIDDEN POPULATIONS

Motivation: Finding and Understanding Hidden Populations

Respondent-driven sampling (RDS) was designed to uncover *hidden* populations, whose membership is unknown and/or unwilling to cooperate, and draw valid statistical inferences about them (Heckathorn, 1997). ISDM uses respondent-driven sampling to guide data collection from a populations whose members do not necessarily have salient differentiating characteristics, and who may not be especially motivated to cooperate with surveyors.

Thus far, RDS has primarily been used to identify and study criminals, drug-users, HIV victims, or the homeless, but the problem of hidden populations is far broader. The decennial census falls short of exhaustiveness on account of multiple missing groups. More broadly, even in a mature democracy such as the U.S., the beliefs and values of 50% of the population are often absent from the public arena because they do not vote or do not show up for jury duty. In meetings of voluntary organizations, including schools and churches, the views of those who sit silently at meetings are hidden, and even more so are the views of those who choose not to attend at all. Those who choose not to engage are sometimes, but not always, silent because they agree with those who do engage. Another possibility is that they intend non-cooperation or even subtle sabotage (e.g., "weapons of the weak" discussed by **Scott, 1985)**. Standard probability sampling methods draw a representative sample of an unrepresentative portion of the target population. They fail to document the fraction of the "silent majority" who choose not to express dissent, but who also intend to be non-cooperative.

Contrast between RDS and Traditional Sampling

Figure 2 illustrates the difference between RDS and traditional sampling and estimation (discussed in Salganik and Heckathorn, 2003). In the traditional approach (Figure 2A), estimation is made directly from the sample to the population; if the sampling process is not random, i.e., if all population members do not have an equal chance of being selected, population estimates made directly from the sample will be biased. The most effective methods for reaching hidden populations are *chain referral* or *snowball* methods, which are far from random (Figure 3B): Because existing respondents refer new respondents, people with large numbers of social connections will be oversampled. Attempts to define artificial frames for random sampling, such as individuals at specific (publicly accessible) sites at specific (safe) times of day, effectively redefine the target population to exclude its hidden portion. Therefore, it has been thought that unbiased population estimates were not possible from samples that effectively find members of hidden populations.

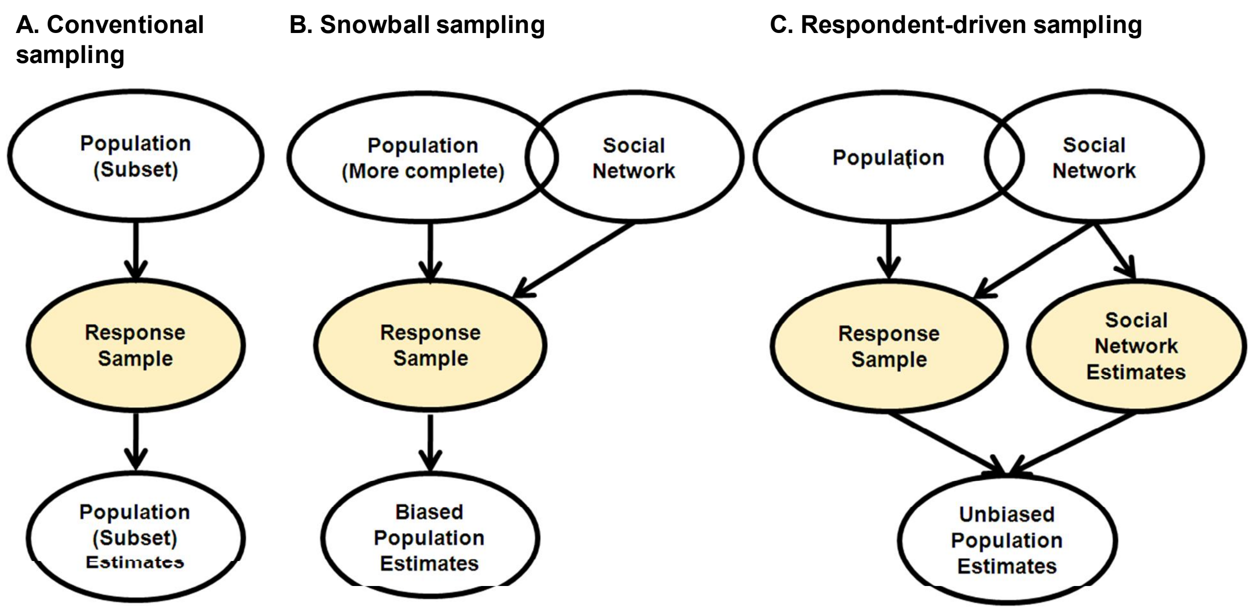

Figure 2. Traditional sampling (left), snowball sampling (middle), and RDS version of snowball sampling (right).

RDS resolves this dilemma (Figure 3C). It meets the double challenge of effectively *reaching* populations that don't want to be sampled and at the same time accurately *inferring* their characteristics without bias. It combines (1) a form of chain referral sampling that identifies members of a hidden population by tracing paths in its social network, and (2) derivation of unbiased conclusions about the hidden population based on the social connectedness of informants. As shown in Figure 2C, population parameters are not directly estimated from sample measures; social network parameters are estimated from the sample and then used to estimate the population parameters. By dealing with the social network explicitly, RDS avoids overestimating the frequency of characteristics associated with more accessible members of the population.

How RDS Works

Like other snowball sampling methods (Figure 3B), the RDS sampling process (Heckathorn, 1997; Salganik & Heckathorn, 2004) begins by selecting and interviewing an initial set of respondents (or "seeds") already known from pre-existing contacts with the subject population. Subsequent sampling is determined by the friendship networks of successive waves of respondents. The data collection portion of RDS (Figure 3C) both incentivizes this process and reduces its intrusiveness. Respondents are paid not only to participate but to recruit additional participants. Each respondent receives a small number of unique, traceable recruitment coupons, which they are asked to give to other people they know in the target population (names are not revealed directly to the surveyors). Respondents receive a second reward when people they have recruited choose to participate (conditional on qualifying as a member of the target population and on the existence of a reciprocal link to the initial respondent). New participants also receive recruitment coupons, and the process continues in successive waves until the desired number of participants is reached.

RDS estimates the portion, PP_A, of the target population with some characteristic, A, by the following formula (Heckathorn, 2002; Salganik & Heckathorn, 2004):

$$\widehat{PP_A} = \frac{\widehat{D_B} \cdot \widehat{C_{B,A}}}{\widehat{D_A} \cdot \widehat{C_{A,B}} + \widehat{D_B} \cdot \widehat{C_{B,A}}}.$$

where D_A is the average degree, i.e., number of social connections, of members of the population with characteristic A; D_B is the average degree of members without A; $C_{A,B}$ is the proportion of social connections belonging to a person with A that cross group boundaries, i.e., link to a person without A; and $C_{B,A}$ is the proportion belonging to a person without A that link to a person with A. When some very general testable assumptions are satisfied, network parameters on the right side of the equation are estimated directly from the sample:

$$\widehat{D_A} = \frac{n_A}{\sum_{i=1}^{n_A} \frac{1}{d_i}}. \qquad \widehat{D_B} = \frac{n_B}{\sum_{i=1}^{n_B} \frac{1}{d_i}}.$$

$$\widehat{C_{A,B}} = \frac{r_{AB}}{r_{AA} + r_{AB}} \qquad \widehat{C_{B,A}} = \frac{r_{BA}}{r_{BB} + r_{BA}}.$$

where n_A is the total number of sampled respondents with characteristic A; d_i is the degree of respondent i; $r_{AB} = r_{BA}$ is the number of social connections involving a member with A and a member without A; and r_{AA} and r_{BB} are the corresponding homogenous associations. Recent work (Heckathorn, 2007) extends RDS analysis to continuous variables and to cases where respondents non-randomly recruit among people linked to them in the social network.

Validity

Small-world results in social network theory (i.e., "six degrees of separation"; Watts, 1999) ensure that everyone *can* be reached by a small number of recruitment waves even in a sparsely connected hidden population. By harnessing intra-group social pressure as an additional inducement to participate, RDS data collection mitigates oversampling of more cooperative members of the population ("volunteers"). It also by-passes the requirement that respondents reveal sensitive information about friends to researchers, because those to whom respondents give coupons make their own choices about participation.

RDS estimates are unbiased if seeds are drawn from the target population with probability proportional to their degree. Even if seeds are selected non-randomly, bias is small, on the order of one divided by the sample size, and asymptotically vanishes as the number of recruitment waves increases. The "law of large numbers for regular Markov chains" shows that as the sample expands wave by wave, the composition of the sample becomes independent of its initial starting points and approximates the population equilibrium at a geometric rate (usually requiring only 4 to 6 recruitment waves in all). Empirical data confirm that RDS sampling fits the path-independence assumptions that define a Markov process. The beauty of this methodology is that whether or not the less visible portion of the population differs from the more visible portion, RDS will effectively reach and accurately represent both.

CULTURAL DATA AND CONSENSUS

Motivation: Identify Distinct Cultural Groups

Snowball sampling will typically span boundaries among groups that differ in beliefs, values, and practices. Members of socially connected populations may still differ significantly in their beliefs, values, and behaviors. Cultural Consensus Analysis (Romney, Weller, and Batchelder, 1986; Batchelder & Romney, 1988; Weller, 2007) is a powerful collection of statistical methods for identifying shared cognitive elements characteristic of a population and/or its subpopulations.

CCA is a break from the tradition in which cultural anthropologists ignored (or took for granted) the empirical distribution of cultural contents. Cultures have been viewed, implicitly or explicitly, as self-contained indivisible wholes that belong to predefined populations. Depending on the specific anthropological school, culture has been defined as a modal personality (Boas, Mead, Benedict), an autonomous symbol system or web of meanings (Geertz), a self-sustaining system of relationships among social roles (Malinowski, Radcliffe-Brown), a culturally consistent mode of thought (Nisbett) or set of value priorities (Hofstede), or adaptations to a common ecology of terrain, climate, and natural resources (Harris, White, Diamond). To describe these unified systems, cultural anthropologists and psychologists freely combined elements that they observed at different times and places or elicited from different individuals in the relevant population. Ethnographic studies of the same "culture" sometimes arrived at significantly different subjective conclusions. Variation was ignored, and culture was effectively treated as a free-floating abstraction.

Two developments in cognitive anthropology broke from this tradition and provide the starting point for Cultural Consensus Analysis. The first development is inspired in part by research on individual difference in cognition and problem solving, which suggests that expertise is based on the accumulation and organization of knowledge in specific domains (e.g., Ericsson & Smith, 1993). At the same time, rather than regarding domain knowledge as an accumulation of discrete facts, expertise is associated with efficient and effective organization of that knowledge for cognitive tasks (e.g., Glaser, 1989; Weiser, Lawrence, & Engel, 1983; Shoenfeld & Herrman, 1982; Larkin, et al., 1980). In response, cognitive anthropologists have undertaken cross-cultural studies of specific domains, such as animals, plants, marriage, healing, collaborative planning, etc. In doing so, they have found diverse domain-specific knowledge structures that support inference – *schemata* (Garro, 2000; Strauss & Quinn, 1997; Shore, 1996; D'Andrade, 1995; D'Andrade & Strauss, 1992; Holland & Quinn, 1987) or *mental models* (Ross, 2004) – which are loosely integrated with one another, if at all.

Second, cognitive anthropologists have supplemented traditional ethnographic data collection methods (e.g., participant observation and semi-structured or unstructured interviews) with more rigorously controlled and standardized elicitation methods borrowed from cognitive psychology (Ross & Medin, 2005; Weller & Romney, 1988). These include: similarity-based sorting, listing category instances, categorizing objects, ranking or rating by specified criteria, evaluating the truth of sentences, filling blanks in sentences, matching items in one list (e.g., possible causes) to items in another (e.g., possible effects), identifying the odd item in a set of three, and estimating the probabilities of alternative conclusions from specified premises (e.g., about categorical or causal relationships). Cognitive anthropology has changed the view of culture to one that is more differentiated (in terms of domain specific beliefs and practices), better grounded in individual cognitive processes and knowledge, and more empirically testable.

Cultural Consensus Analysis fills a crucial gap between eliciting cognitive contents or processes of *individuals* and attributing cognitive contents or processes to *populations*. CCA identifies the pattern of variation within a population and uses it as a basis for statistical inferences about the attribution of shared content. More specifically, it assesses the degree to which cultural elements are shared within a group, the systematic variation in content among subgroups, and the degree of representativeness (or "cultural competence") of individuals with respect to that content.

How CCA Works

In traditional psychometric test theory, as shown in Figure 3A), an individual's quality of *performance* on a set of questions is assessed by comparing that individual's *responses* with a pre-specified *answer key*. Cultural Consensus Analysis, by contrast, is designed for an outsider who poses questions without knowing the correct answers, without knowing who might answer most reliably, and without in fact knowing if correct answers exist for the group in question. CCA identifies culturally correct sets of answers and the most trustworthy respondents even when informants within the same group do not agree.

Cultural Consensus Analysis adds three main steps to standard population estimation, which can be traced by reference to Figure 3B: (1) It computes the *agreement* of each individual respondent with the mean of other respondents across questions. (2a) It uses average agreement across respondents (i.e., degree of population consensus) to evaluate the assumption that an *answer key* exists with respect to the questions. If not, ways of subdividing the population into culturally coherent subpopulations are explored. Population estimates and agreement within each subpopulation .are then calculated for each subpopulation separately. (2b) If the hypothesis of a single answer key is not rejected for a population or subpopulation, degree of agreement with the consensus estimate is used as an estimate of an individual's *competence* in the domain. (3) The consensus estimate is then recomputed by weighting each individual's response by that individual's competence. The process is iterated until estimates are stable. We will briefly mention some salient points regarding each of these steps.

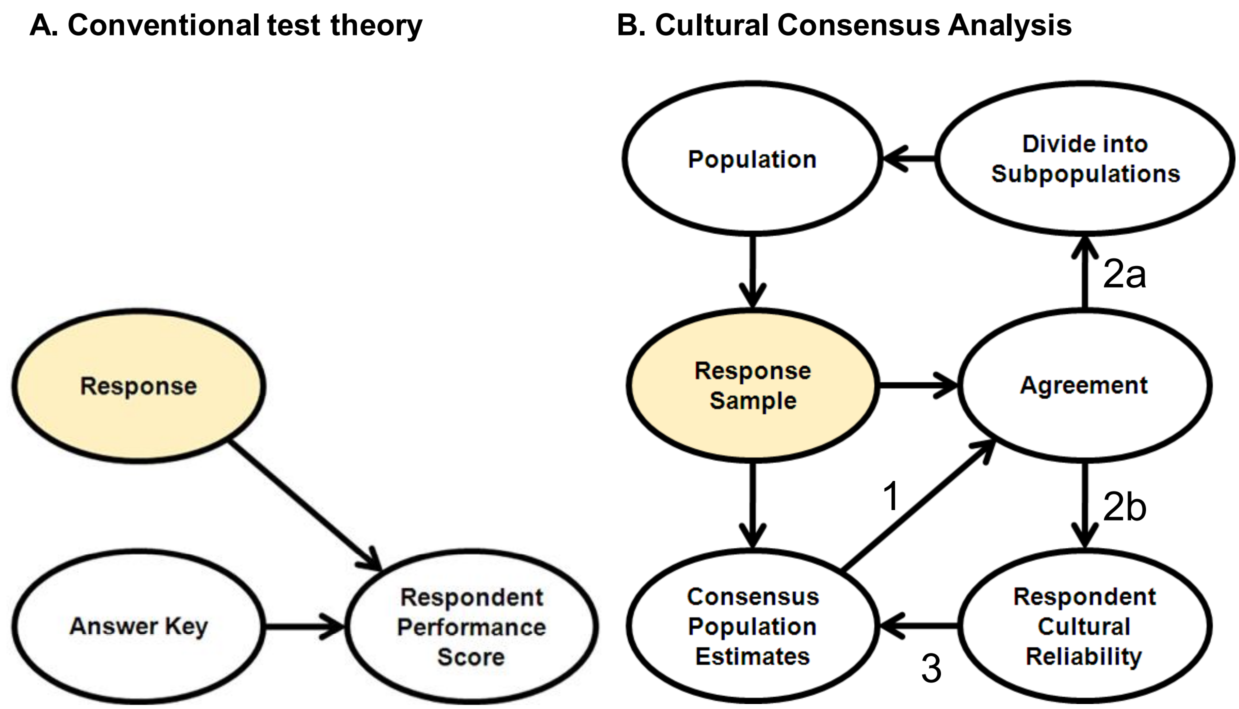

Figure 3. Traditional Test Theory (left) and Consensus Theory (right).

A distinctive aspect of CCA compared to other information aggregation methods is that it does not pool answers to one question at a time; it analyzes the entire pattern of responses to a *battery* of questions on the same topic, in order to infer *domain* competence. If questions are categorical CCA assumes that if the respondent *knows* the answer to a question, he answers it correctly; if not, he guesses randomly (in the simplest version). Thus, a pair of respondents must agree on the answer to a question when they both *know* the right answer, but otherwise will agree only by chance. The expected proportion of matching answers for respondents i and j (M_{ij}) depends on their respective levels of domain competences (D_i and D_j), i.e., their probability of knowing the correct answer to questions on the topic. The theory estimates the probability that informants i and j both *know* the right answer as a linear function of the observed proportion of matches (M_{ij}), corrected for guessing. Individual levels of competence (e.g., D_i and D_j) are then estimated from the *informant-by-informant* matrix (D_i D_j) by *minimum residual factor analysis*, which estimates the diagonal cells (D_i^2) by minimizing the squared errors of off-diagonal predictions.

If cultural knowledge is elicited by non-categorical questions (e.g., rank ordering items, or rating) or by open-ended questions with multiple answers (e.g., free listing), the formal theory of question answering does not apply, and a factor analysis of the matrix of observed matches is used without correction for guessing.

The first factor extracted from the factor analysis corresponds to an “ideal informant.” If this factor accounts for most of the variance in the *informant-by-informant* matrix, the assumption that a single answer key exists is supported), and variations in respondent answers are legitimately explained in terms of competence rather than cultural differences (step 2b). If the first factor leaves significant variance unaccounted for, there may be no cultural model at all, or else divergent cultural models may be associated with subsets of the original population (step 2a) (e.g., Mueller, 2008; Smith & Batchelder, 2008; Handwerker, 2008; Atran, Medin, & Ross, 2005; Grant & Miller, 2004). CCA serves as an exploratory tool for finding coherent subpopulations and their associated models, based on indicators such as: negative or highly variable loadings for individual respondents on the first factor, between-group variance accounted for by the second factor, and/or deviations from expected agreement accounted for by demographic factors like socio-economic class, age, gender, generation, ethnicity, profession, or social network neighborhood (Ross, 2004; Atran, et al. 2005). Proceeding in the opposite direction, CCA can also be used to identify shared elements across populations that are conventionally regarded as distinct (Atran et al., 2005).

The output of the third step is a posterior probability for the correctness of each possible answer to each question, within a sufficiently coherent population. Weighting by respondent competence produces a more accurate result than majority rule (for categorical responses) or averaging (for continuous quantities), especially when there are small numbers of sources who vary in competence. CCA permits calculation beforehand of the minimal number of respondents needed to reconstruct the correct answers with a desired level of accuracy as a function of a given level of agreement in the group. RDS estimates the number of required sampling waves necessary to achieve the desired effective sample size.

Validity

In sum, Cultural Consensus Analysis provides an information-pooling methodology that does not incorporate a researcher's prior beliefs about the correct answers, the existence of subgroups, or the calibration of respondents. The model is applied on the basis of three testable assumptions: *common truth* (i.e., there actually exists a cultural model with correct answers applicable to all members of the specified group), *local independence* (i.e., respondents' answers are independent except for their shared knowledge of the true answer), and *monotonicity of items* (i.e., respondents who have more competence on any subset of questions will have more competence on all subsets). The latter assumption requires that questions reflect the same body of knowledge, i.e., cultural knowledge has been carved up by domain. Aggregation is based on the principle that agreement is an indicator of accuracy among respondents who are attempting to describe the same common truth.

Agreement with other respondents has been empirically validated as a measure of cultural competence in studies showing a high correlation between an individual's (agreement-based) competence and the individual's test-retest reliability (**Boster, 1985**) as well as the individual's internal logical consistency, e.g., avoidance of intransitive judgments (**Weller, 1984**; Brewer et al., 1991); moreover, in cross-domain comparisons, informants are more logically consistent in domains that have high overall consensus (**Brewer et al., 1991**). These findings support CCA's assumption that errors are due to lack of knowledge: When informants do not know the answer, they guess, leading to inconsistencies across an individual's responses to the same or to logically related questions. Brewer et al. (1991) found that cultural competence in one domain is not correlated with competence in another domain, supporting the domain-specific characterization of the cultural knowledge. Finally, simulation studies show that CCA accurately recovers correct answers from a small number of fallible respondents (Karabatsos & Batchelder, 2003).

Social network analysis and Cultural Consensus Analysis provide complementary measures to enhance planning of operations intended to influence a population. Influence operations will be more effective if they target attitudes that are representative of a culturally coherent *and* socially interconnected target group. Chances of success are also improved if operations enlist the cooperation of *individuals* in those subgroups who are culturally competent *and* central in networks of friendship, trust, information, or authority (Borgatti, 2005). Since diminishing interconnectedness tends to parallel decreasing levels of agreement as group size increases and interaction decreases (e.g., neighborhoods, districts, cities, respectively), the methods can be used in combination to select optimal ways of subdividing larger populations for purposes of persuasion.

INFLUENCE DIAGRAMS FOR CULTURAL MODELING

Motivation: Represent Inferential Relationships for Dynamic Reasoning

The output of Cultural Consensus Analysis is a set of *discrete* answers to a battery of questions together with a measure of confidence, or posterior probability, for each. The only connection among the questions is the assumptions that they belong to the same substantive domain, which gives rise to a single competency and justifies the calculation of respondent reliability. There is no explicit indication of how multiple answers interact with one another in the group's reasoning about the relevant domain, i.e., how some answers are used in combination to derive others in dynamic cognitive tasks, such as situation assessment and decision making. In other words, there is no attention to structure within the set of consensual answers, manifested in covariation of answers across time or situations.

Cognitive research on individual differences in problem solving skills shows that expertise is not simply the accumulation of knowledge, but also the efficient organization of that knowledge for explanation, prediction, and choice (e.g., Glaser, 1989; Weiser, Wolfe, Lawrence, & Engel, 1983; Shoenfeld & Herrman, 1982; Larkin, et al., 1980). Cultural models are not compendia of answers to discrete questions, as provided by conventional applications of CCA. They are "theories" of the domain that are used to generate new beliefs and decisions in a dynamic environment (Atran, Medin, & Ross, 2005). This theoretical structure is embodied in relatively stable inferential relationships (Sieck & Rasmussen, 2008; Garro, 2000; D'Andrade, 1995), rather than answers to questions about discrete beliefs, values, or intentions.

Similarly, cognitive anthropologists have explained cultural influence and change as active accommodation to new ideas and behavior, not as a simple replication or imitation (Sperber, 1996; Atran, Medin, & Ross, 2005). Understanding of stable, shared inferential relationships is necessary to support prediction of potential or actual *change* in views or actions over time. For example, different subpopulations, with similar cultural theories, may be exposed to different exogenous factors – e.g., climate, invasion, immigration, or peaceful spread of ideas from neighboring cultures or embedded subcultures (Ross & Medin, 2005; Reyes-Garcia et al., 2003). These changes may lead to indirect effects on *other* beliefs, values, or practices, via relatively more stable inferential relationships. Thus, systematic patterns of covariation among consensual beliefs, values, and choices across groups, or within the same group across different time periods, may arise if they are applying the same inferential relationships to different "premises," i.e., exogenous conditions or influences (Atran, Medin, & Ross, 2005).

Note that change in consensual answers over time, and interdependence of consensual answers across subpopulations, do not violate CCA assumptions, which require independence of *responses* across questions and respondents *given* the current set of consensual answers for a particular coherent subpopulation. Although they do not rule it out, CCA methods do not discern interdependent change because: (i) culturally correct answers by definition cannot vary *within* a culturally coherent group; (ii) CCA does not look at covariation of consensual answers *across* groups, or within the same group across time; and (iii) CCA does not distinguish (a) relatively stable inferential relationships, (b) dynamically changing environmental factors and influences, and (c) dynamically changing concussions drawn by applying the inferential relationships to the changing inputs.

CCA needs to be supplemented by a systematic approach to elicitation of inferential relationships such as causal mechanisms and category membership from individual respondents, inclusion of those relationships in CCA analysis, and the composition of the relationships into *cultural models*, i.e., active computational structures for interpretation, explanation, prediction, and decision making that can be sensibly attributed to culturally coherent groups. Such models are particularly important when the results will be used to *influence* subject populations.

Although cognitive psychologists have demonstrated the importance of knowledge structures, they have nor settled on a consistent, rigorous format for representing them. Cognitive anthropologists refer to *schema theory* (e.g., Garro, 2000; Strauss & Quinn, 1997; Shore, 1996; D'Andrade, 1995), but schemata are usually defined in functional terms, as active processes of interpretation: Triggered and instantiated by contextual cues, they fill gaps in observations; activate explanatory assumptions, expectations, and goals; and prime appropriate actions. *How* they accomplish these cognitive tasks is often not specified, and at other times is restricted arbitrarily to fit the theoretical predilections of the researcher. Suggested formats and computational mechanisms include if-then (production) rules, inheritance in semantic networks, frames, similarity-matching to prototypes, goal-method hierarchies, qualitative mental models, and spreading activation in a neural net architecture.

ISDM represents cultural models of relevant populations as *Bayesian influence diagrams*. Unlike schema theory, influence diagrams provide a formally coherent, deeply studied, widely used, fully specified calculus for reasoning under conditions of uncertainty and multiple objectives (Pearl, 1989, 2000; Lipshitz & Cohen, 2005). Influence diagrams are both precise enough and general enough to account for reasoning in a variety of contexts. Moreover, they have a unique pragmatic advantage: they can be embedded in *decision models* for planners, to support course of action evaluation in operations intended to influence populations. Effective planning of such operations requires a *dynamic* and *systemic* view of cultural contents. Planners need to determine which beliefs and values offer the most leverage over desired target population behaviors, and the kinds of information or actions that might induce the required changes.

How Influence Diagrams Work

Decision analysis is the practical implementation of ideas developed by economists, logicians, and statisticians under the rubric of *decision theory*. Traditionally, decision analysis has been associated with a set of related by discrete modeling paradigms: decision trees for choice with uncertain outcomes, event and probability trees for uncertain inference (explanation, prediction, and evidential updating), and goal hierarchies for choice based on multiple criteria (Raiffa, 1968; Keeney & Raiffa, 1976; von Winterfeldt & Edwards, 1986; Watson & Buede, 1987). Unless extensively pruned, trees display every possible combination of variable states as a distinct path (see Figure 4A and Figure 5A), and require probability assessments for every branch conditional on the entire path leading up to it. Despite the advantage of concreteness, the exponential growth in model size with number of variables imposes severe burdens on visual intelligibility, assessment, and computation even with relatively small models.

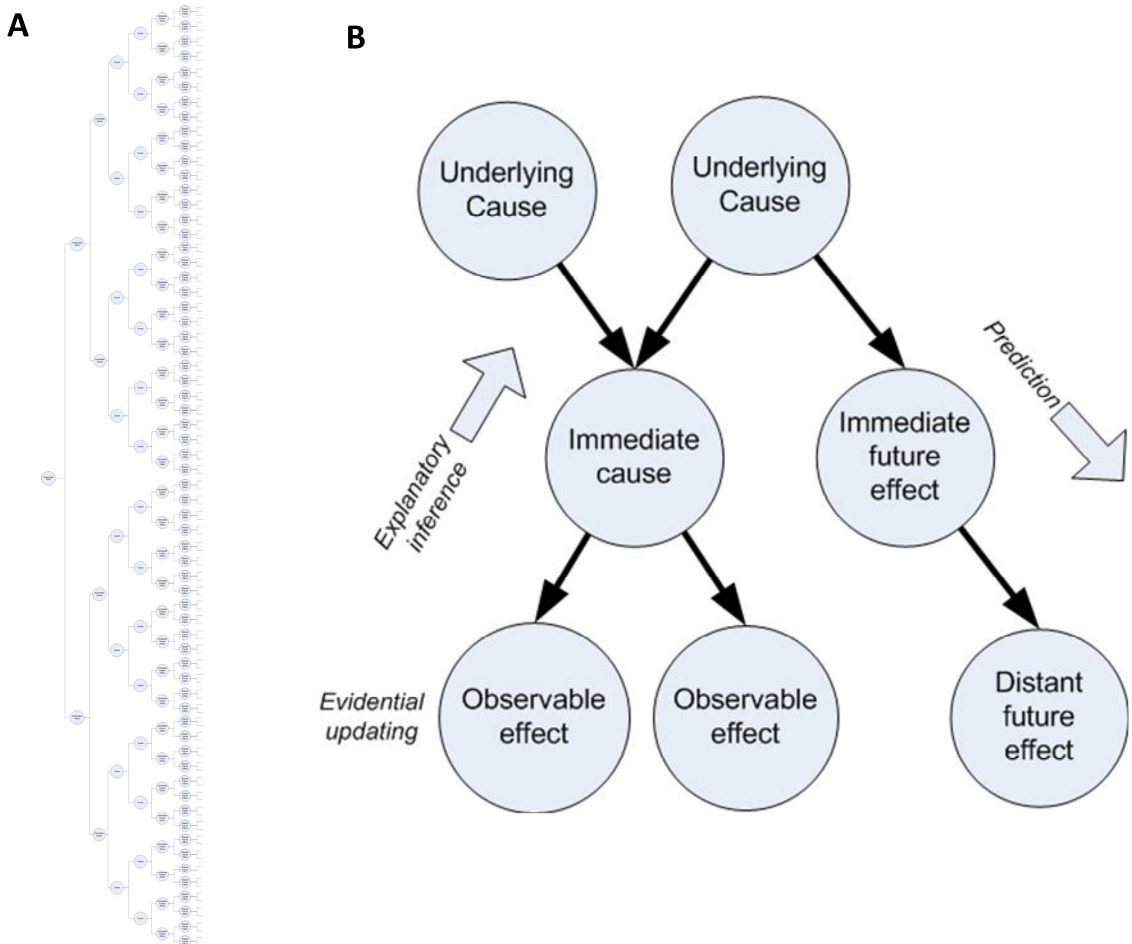

Figure 4. A is the full event tree for seven binary variables, showing every possible combination of events as a path ($2^7 =$ 128 paths). B is the corresponding Bayes net, which shows causal relationships among the seven variables. (Variable names are descriptive and have no relation to functionality.)

The most exciting work in decision analysis over the past two decades applies a more powerful modeling technology to these problems, i.e., *Bayesian networks* and *influence diagrams* (Howard & Matheson, 1989; Pearl, 1988). As illustrated in Figure 4B, Bayes nets provide a single integrated framework for reasoning about uncertainty, including updating hypotheses in light of new evidence (Edwards, 1968), hierarchical explanatory inference (Schum, 1994), and forecasting (MacGregor, 2001). Each node in a Bayes net represents a variable whose states are events or conditions. Arcs represent direct causal influence exerted by parents on the child. The resulting diagram is far more compact than the corresponding probability tree (Figure 4A), and supports visualization and assessment by a natural causal representation.

Influence diagrams exploit independence among variables to reduce the assessment burden. Prior probabilities are assessed for the parentless causal nodes in Figure 4B. Conditional probabilities are assessed for each child node conditional only on combinations of states of its parents. The number of parameters grows exponentially with the number of local parents, therefore, not with the total number of variables in the model. The probability tree in Figure 4A requires 127 independent assessments, while the influence diagram in Figure 4B requires only 14, and the advantage grows with the size of the model. Moreover, *canonical operators* (e.g., causal, probabilistic versions of logical operators, called Noisy-*Or* and *Noisy-And*) have been developed both to clarify the semantics of the probabilities and to reduce the number of assessments to a manageable number, which is *linear* in the number of parents (Henrion, 1989: Heckerman & Breese, 1996; Diez & Druzdzel, 2005). In addition, influence diagrams are more flexible and general than trees. Information about known states of any subset of variables can be set and automatically propagated across the network, generating posterior probabilities for all other variables conditional on that information. A probability tree must be rearranged for each new specification of factors as knowns and unknowns.

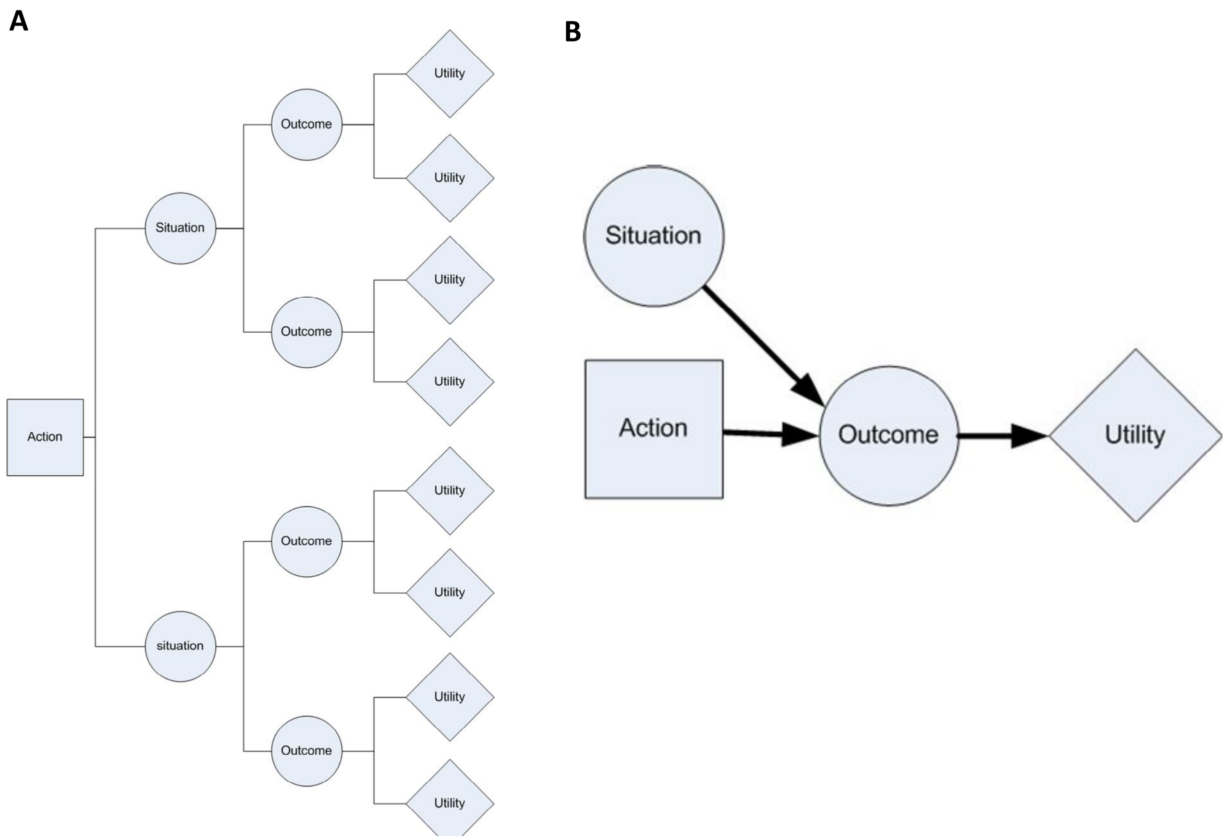

Figure 5. A is the decision tree for a simple choice with uncertain consequences. B is the corresponding influence diagram.

Influence Diagrams are *decision models* that incorporate Bayes nets and in addition support choice under uncertainty and tradeoffs among multiple objectives (Owen, 1989; Clemen & Reilly, 2001). A maximally simple decision problem (Savage, 1954; Raiffa, 1968) involves three elements: uncertain conditions (as in Bayes nets), decisions, and consequences, as illustrated in Figure 5. Rectangles are decision nodes, representing variables whose states are actions under the control of the decision maker. Ovals are chance nodes, which represent uncertain conditions with the same parameters as Bayes nets (Figure 4B): conditional probabilities must be assessed for states of chance nodes conditional on all combinations of states of their parents (which may include both decisions and other chance nodes); prior probabilities must be assessed for states of a parentless chance node. (For convenience, we will sometimes call uncertain conditions *outcomes* when they are influenced by a decision and *situation* factors when they are not, as in Figure 5. However, the outcome of one action may be a situation factor with respect to a subsequent action, as will be seen in Figure 7.) Diamonds are utility nodes, representing consequences in the form of a utility assignment, or degree of preference, to every combination of their parents' states. An action is optimal if it maximizes subjectively expected utility (SEU), which is the sum over consequences of utility multiplied by probability of the consequence given the action.

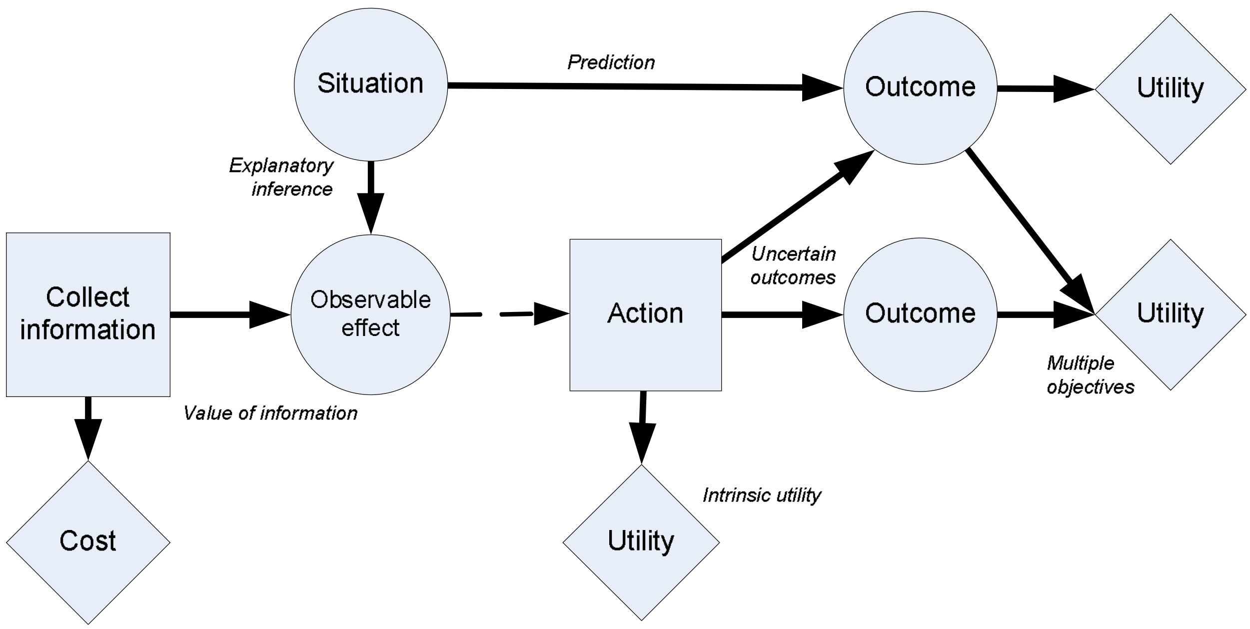

Figure 6. A more complex influence diagram.

Figure 6 illustrates a somewhat more complex decision problem. It combines choice under uncertainty, as in Figure 4B, with *multiattribute utility* decision making (Keeney & Raiffa, 1976) and *value of information* analysis (LaValle, 1980, 1968; Cohen & Freeling, 1980).

Multiattribute utility. Multiple utility nodes in Figure 6 represent different objectives, or sources of value. Actions can have multiple outcomes, and outcomes may influence the achievement of different objectives – leading to difficult tradeoffs. Parameters for the model in Figure 6 include importance weights for each objective and scores for each outcome on each objective that it influences. An action may also have intrinsic utility, in addition to instrumental utility as means to desirable outcomes, Intrinsic utility is represented in Figure 6 by a direct link between the decision node and utility. Total utility is the sum of utility from all different sources, weighted by importance.

Value of information. The observable effect of a chance node often provides *evidence* regarding its state. Evidence may be produced by the chance node in conjunction with optional information collection actions, as in Figure 6. The decision maker therefore faces a secondary decision: whether or not to obtain that evidence in order to reduce uncertainty about the situation. Learning about the situation will also reduce uncertainty about decision outcomes that result jointly from the situation and the action. In sum, because the uncertain situation influences not only the future outcome but also potentially observable present effects, the latter can serve as *evidence* regarding aspects of the current situation that explain the evidence and help predict the future.

Because collecting the information incurs some cost (i.e., negative utility), doing so is worthwhile only if those costs are outweighed by the benefit. This benefit, i.e., *value of information* (VOI), is the difference between the expected utility of the decision now and the expected utility of the decision with the additional information. Value of information is zero unless there is a possible observation that would change the decision, leading to higher expected utility than the currently preferred option. VOI is proportional to the probability that the evidence will change the decision and the difference in expected utility that results. A dotted arc from *observable effect* into the decision node means that the state of the uncertain condition *will* be known to the decision maker before the decision is made. If the decision maker chooses not to obtain the information, the state of the evidence is "not available"; if he does choose to collect it, it is correlated with the state of the situation (Shacter, 2007).

The collection decision in Figure 6 has information as its sole purpose, but *any* action can produce information that is useful for subsequent decisions; any outcome of any action may serve as an informational input to subsequent decisions. In Figure 7, the situation at any given time is the result of previous action; observable effects enable the decision maker to react to the situation as it exists and to forecast the situation that will result from a new action. A *course of action* is a total strategy: a choice of one state from each decision variable in the model, contingent on states of uncertain variables that will be known when the decision is made. Figure 7 also illustrates how feedback is modeled without the use of explicit directed loops in an influence diagram; the same condition at different points in time is represented by different variables (or, equivalently, the same condition indexed by time). A condition *at a given time* may inform an action, which in turn affects that condition *at a subsequent time*.

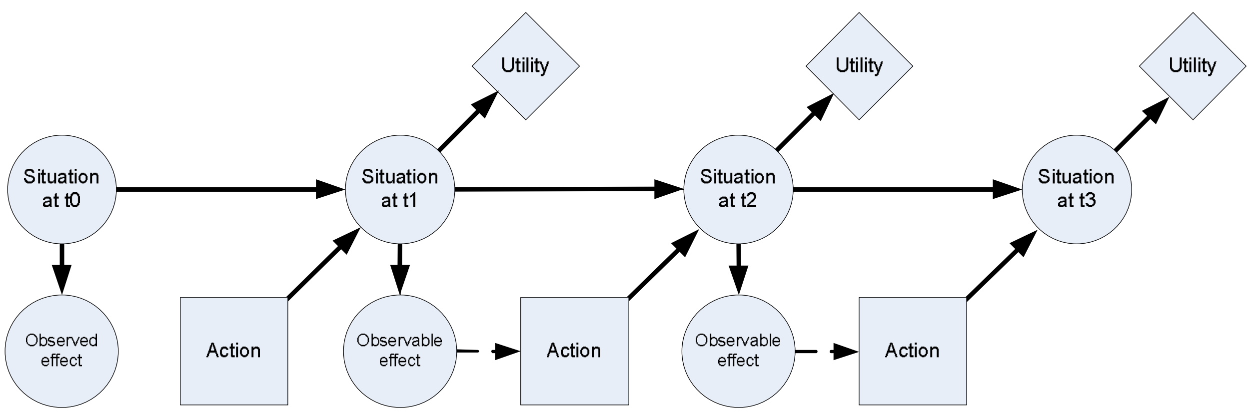

Figure 7. Decisions influence an evolving situation state based on its observable effects.

An advantage of influence diagrams over traditional decision trees is that they permit more explicit analysis of the causal interactions underlying outcomes in a given situation. Richer situation models stimulate *option generation*, rather than simply supporting selection from a predetermined set of options. In particular, any uncertain condition that has a utility node among its descendents is a potential lever for improving consequences. The *value of control* (VOC) for a variable is the gain in expected utility that would result from forcing that variable into its most advantageous state. For example, in Figure 8 each situation is a function of the previous situation state and action, as well as a specific immediate condition. The decision maker might improve expected utility by finding a way to control the immediate conditions of action. An even more efficient way to improve expected utility, if feasible, might be to influence their common underlying cause. (If a variable such as Immediate condition at t2 is controlled, it is no longer evidence for the operation of the processes that normally determine it, viz., Underlying cause. Therefore, to calculate VOC, all causal inputs to the variable are severed before setting the desired state. Because it filters out incorrect explanatory inferences, VOC is a more valid indicator of a factor's importance than simply varying its state and observing the impact on expected utility.)

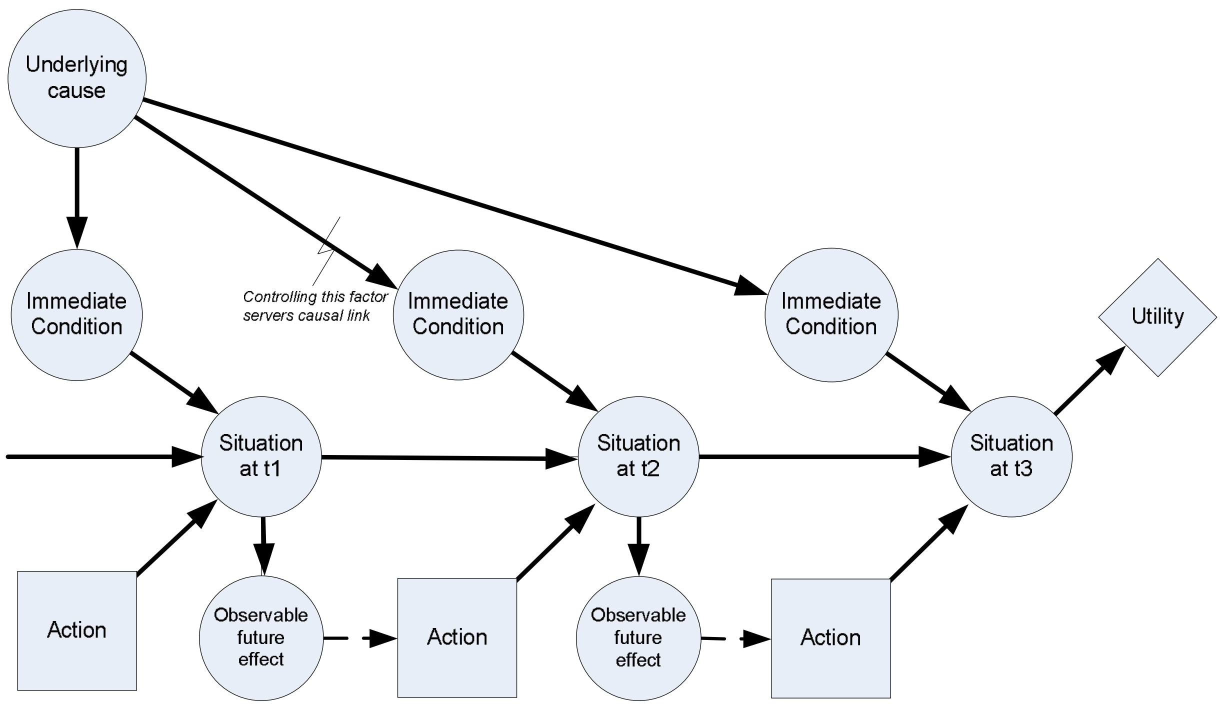

Figure 8. Example of value of control.

Richer situation models also stimulate generation of options for information collection. An information collection option need not be represented explicitly, as it was in Figure 6. Any uncertain condition that has a diagnostic and/or causal relationship to the utility of a subsequent decision is a potential source of evidence about outcomes of that decision. VOI in this mode is a form of sensitivity analysis, which captures the practical impact of information. (The variable for which VOI is calculated cannot itself be an outcome of the decision to be supported, on pain of violating causal and temporal constraints. This requirement can be sidestepped by locating uncertainty in the situation. Outcomes can be represented as deterministic functions of actions, uncertain conditions, and uncertain *rules* of the form, *If <combination of states of conditions and actions>, then <outcome state >*. Conditional probabilities of outcomes given combinations of states of their parents may now be represented as probabilities of the truth of these rules. The value of observing the state of a situation factor-rule can be calculated because it is not influenced by the decision. This approach is simplified if relationships between outcomes and parents are represented by means of *canonical operators* such as *Noisy-Or* and *Noisy-And* (Henrion, 1989: Heckerman & Breese, 1996; Diez & Druzdzel, 2005). In that case, there is one independent parameter for each binary parent of the outcome; each parent-child relation is associated with one rule of the form, *If <parent>, then < outcome>*. In this case, VOI is computed for *impact weights*, which can be displayed on the causal arcs rather than as visually distinct nodes.)

In sum, Bayes nets and influence diagrams seamlessly integrate cognitive tasks previously associated with specialized representation formats and computations, and they are solved by increasingly efficient computational algorithms (Shacter, 1986; Pearl, 1989; Shenoy, 1992; Jensen et al., 1994, 2001). Taken together, they help users (1) visualize causal and logical relationships among variables representing actions, uncertain conditions, and objectives; (2) calculate the implications of information about any subset of variables for all the others, (3) identify the optimal combination of choices under foreseeable contingencies, and (4) calculate the pragmatic benefits of collecting additional information or controlling elements of the situation.

Validity

The development of decision theory as an overarching framework for Bayesian probability and rational choice is regarded as among the most significant accomplishments in logic and statistics in the second half of the twentieth century. Decision theory (Savage, 1954; de Finetti, 1964; Jeffrey, 1965; Luce & Raiffa, 1957) formalizes choice in terms of consistency among beliefs, desires and choices. Its normative force is based on the plausibility of the axioms that define consistency, e.g., completeness and transitivity of preferences, and consistency across supposedly irrelevant variations in context (Savage, 1954). It can be proven that an agent's beliefs, desires, and choices satisfy the axioms if and only if the agent chooses actions that maximize *subjectively expected utility* (SEU) – i.e., the sum, over an action's possible consequences, of the utility for each consequence weighted by its probability of occurrence given the action (Jeffrey, 1965).

Decision analysis defines a process and a set of tools for ensuring that beliefs, desires, and choices in a particular problem satisfy the formal decision theoretic constraints (Raiffa, 1968; Keeney & Raiffa, 1976: Brown & Paschoud, 2005; Edwards, Miles, & von Winterfeldt, 2007). More pragmatically, decision analytic tools organize knowledge systematically in terms of decision relevance, provide an explicit shareable model of the evolving problem, support a deliberative process with high face validity, and serve as an auditable record or decision rationale (Lipshitz & Cohen, 2006; Cohen & Lipshitz, 2010).

Cultural modeling does not assume that members of the populations being modeled engage in explicit deliberation, or even that their behavior satisfies axioms of rational choice (*as if* they had deliberated). A Bayes net is a model that *describes* the decisions made by the subject population from an external, third-person perspective. The only assumptions are that subjects *can* be characterized as having beliefs, values, and actions, and that there are causal relationships of *some* kind among them. Nevertheless, a rational model from the perspective of the subject population can be heuristically useful as a way of entering into the subject's point of view, ensuring that relevant factors are identified, assumptions made explicit, and inconsistencies clarified. It can then be converted into a third-person Bayes net. The latter can be incorporated within a *normative* decision model for the operational planner.

CULTURALLY AWARE PLANNING

The planning module of ISDM supports decision making by operational planners who want to understand, anticipate, and, perhaps, modify the beliefs, values, and intentions of relatively unknown populations. Beliefs, values, and actions of populations of interest thus play several roles in ISDM planning models: (1) They are uncertain situation factors that influence outcomes of actions – for example, when the degree of cooperation by the population in providing intelligence will influence the success of a military operation. Such conditions must be understood and predicted in order to determine if the planned primary actions will work. (2) *Changing* a population's beliefs, values, and intentions may be an objective, in which case beliefs, values, or intentions are among the uncertain consequences that determine utility. Influencing attitudes may be a fundamental long-term mission goal (e.g., improved respect for the legitimacy of the host nation government; reduced sectarian conflict), or it may be undertaken in support of actions with other primary objectives, e.g., to encourage cooperation in providing intelligence (see case 1). Changing attitudes almost always implies that *current* beliefs, values, and practices serve as situation factors, as in case 1; that is, current attitudes should be understood before attempts are made to change them. (3) Unintended negative reactions can always occur as by-products of actions undertaken for other reasons, and they can diminish or even outweigh the intended benefits of the actions (e.g., cordoning neighborhoods to search for insurgents, attacking terrorists in mosques, or building co-ed schools). Another reason for actions undertaken to influence attitudes (case 2) is to head off or neutralize such potential effects. Of course, these three roles are not mutually exclusive.

There are very few situations where exact predictions of human behavior can be made with confidence; cross-cultural predictions are even more difficult. The incorporation of cultural knowledge into decision models may sometimes help reduce uncertainty, by narrowing down the ways other groups are likely to interpret a situation and make choices. But cultural knowledge can also help by increasing uncertainty, e.g., if it directs attention to previously unconsidered culturally relevant factors or outcomes. The most important use of decision models may as a qualitative tools for identifying possibilities and developing more robust plans.

ISDM supports the following steps in cultural modeling and planning:

1. **Define the problem** by means of a simple *initial decision model*.
2. **Develop a generic cultural model** with enough detail to cover major factors affecting outcomes of the initial decision model. Prestored templates support this process. Templates have open slots that may be filled by lists from a Catalog of relevant categories and objects. Templates self-aggregate based on shared factors.
3. **Identify information requirements and develop questions.** Prioritize factors in the model in terms of their potential impact on decision outcomes. Expand the cultural model until measurable factors are identified for high priority factors. Compose survey questions for measurable factors. Define relationship of answers to model parameters.

   (Administer the survey with the help of Respondent Driven Sampling (RDS); use Cultural Consensus Analysis (CCA) to distinguish sub-populations and identify prototypical responses for each population.)

4. **Define distinct cultural models** for each population identified by Cultural Consensus Analysis. The models include interventions intended to influence key factors for each population.
5. **Integrate cultural models into decision model.** A final decision model is created by combining the distinct cultural models with the initial decision model. The decision model ranks all interventions – either by population or over all populations – and identifies the most promising combination of actions to achieve desired outcomes.

We will briefly discuss each of these steps in turn, by reference to a concrete example.

Define the Problem

To define the problem, the decision maker specifies an extremely simple decision model. A user-friendly dialog requests four basic ingredients: (a) a decision among specified alternative actions, (b) one or more objectives to be achieved by the chosen action, (c) uncertain conditions upon which success depends, and (d) one or more sources of value or utility (which may be identical with b). In problems of interest to ISDM users, uncertain conditions and objectives typically involve population behaviors. The model is not intended to capture all aspects of current operations. Its purpose is to focus cultural modeling and planning on *estimating* and *improving* the chance that the specified objective(s) will be achieved..

In Figure 9, for example, the question is whether or not to promote the formation of local civic councils in small towns and villages in a host nation – the kind of decision frequently faced by military or civilian planners engaged in Peace and Stability Operations (so-called nation building). The source of value is a desired mission end state, reduced sectarian strife (measured by the frequency of incidents of various types). Uncertainty about achievement of the objective is localized in explicit uncertain situation factors. Achievement of the objective depends on the decision to form the councils and on how members of relevant populations would respond if invited to participate in them. The *And* operator means that the objective is achieved to the extent that members of all population groups would participate.

Users can fill open object places, or *slots*, in factors by dragging and dropping object names from *Catalog* displays provided to support planning. Catalog displays represent hierarchies of object types such as Agents, Times, Places, and Equipment. Each display has an expandable outline structure, in which more general types include more specific subtypes or primitive objects, as illustrated in Figure 10A (left) . Users can create multiple factors at once by dragging and dropping selected folders or objects, at any level, from a Catalog into slots in the Model Composer workspace. Figure 10B shows the result of dragging and dropping four population names into the open {Population} slot in Figure 9.

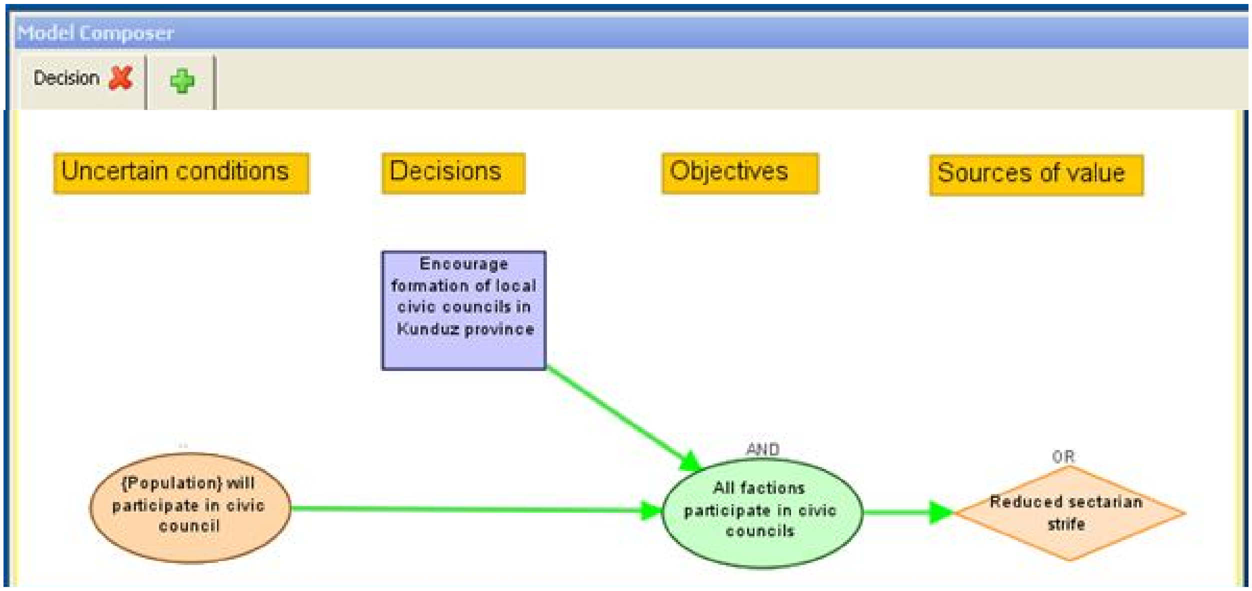

Figure 9. Illustration of initial decision model for problem of whether or not to support formation of local civic councils, in order to reduce sectarian strife.

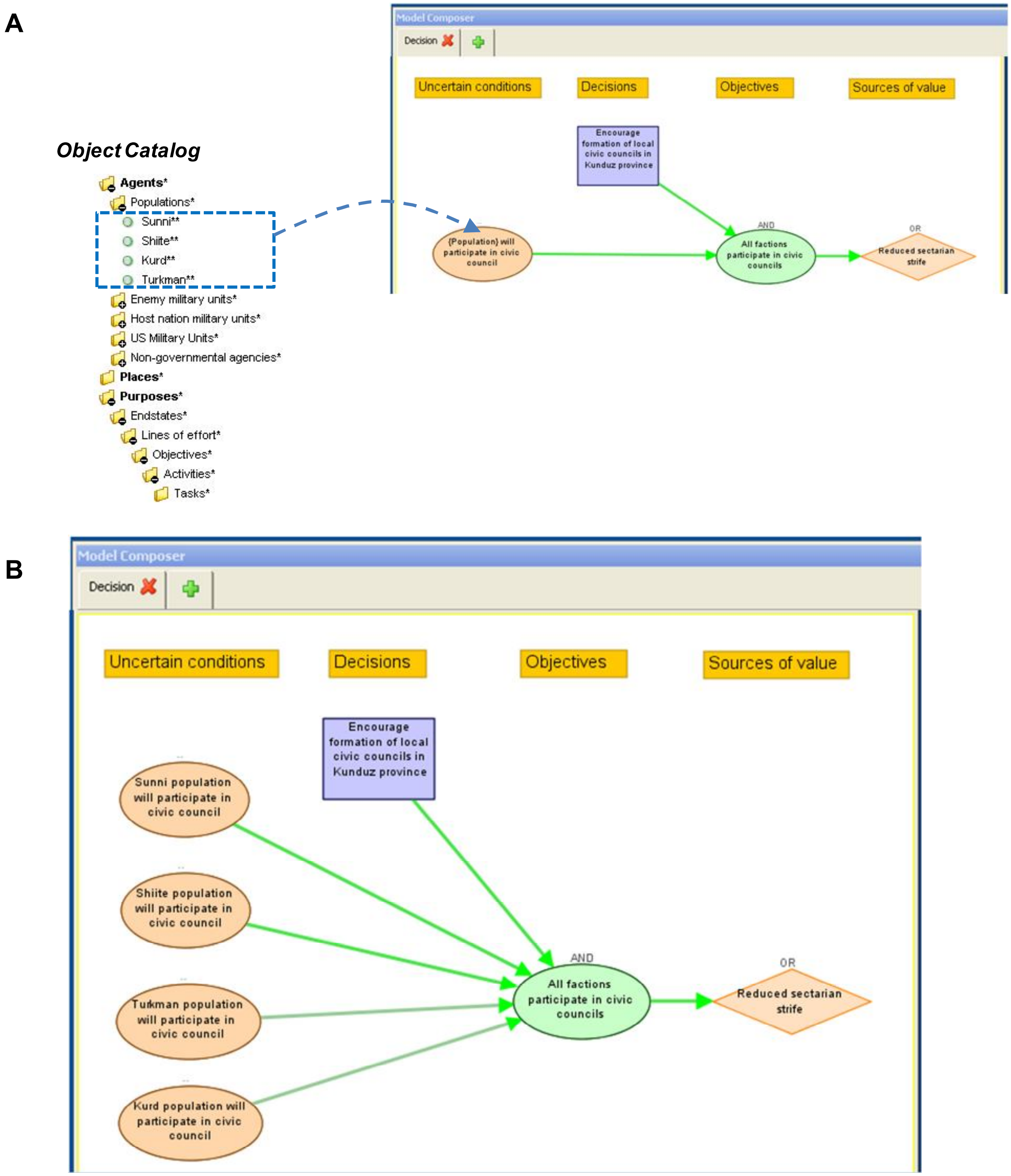

Figure 10. (A) User drags and drops list of populations from a Catalog display into the open slot in Figure 9. (B) Decision model expands to include a factor for each population.

In Figure 10, the user has chosen to specify relevant populations directly. Users may prefer not to prejudge this issue. If so, they can instruct ISDM to let Cultural Consensus Analysis fill in a {Population} slot automatically.

In addition to filling slots from object Catalogs or by CCA, a variety of methods are available to help users define the problem, as a fully integrated part of the planning process itself.

- The user can drag and drop shapes into the Model Composer workspace from a predefined palette, edit and connect them and assess parameters.
- The user can extract the relevant part of a larger *Mission model,* if one exists. (Establishing a successful civic council may be only one of several objectives in a larger line of effort addressing sectarian strife, which is itself only one component of a larger mission that includes, for example, improved governance and economic development.) Parts of a larger model can be copied into separate tabbed Model composer screens manipulated separately, and then re-linked to the larger model..
- *Plan panels* allow users to create tables in which objects from various (sub)types become associated with one another. For example:

  Purposes may be associated with start and end times and locations.

  Populations may be assigned to friendly units.

  Friendly units may be associated with actions/tasks, times, and locations

  Enemy units may be associated with times and locations where unit was observed.

  Each table is a relation (in the database sense), whose rows may be used as factors or events in the decision model. Users can add factors to a decision model by dragging and dropping rows from Plan panels into the Model composer workspace.
- *Specialized displays* are used when it is necessary to hard wire a customized type of representation, such as Timeline, Map, and Decision Model. Events in any one of these displays may be dragged and dropped into any other.
- A store of prepared *Model Templates* provides important guidance, as described in the next section.

Develop a Generic Cultural Model

The next step is to elaborate the initial problem definition model by identifying cultural factors that significantly influence success. Although ISDM provides a general purpose decision analytic modeling capability, its primary purpose is to support relatively non-technical planners, or members of a planning staff, in relatively time constrained environments. Typical users cannot be expected to build models entirely from scratch. For this reason, ISDM utilizes a library of prepared templates customized to a specific problem domain. Templates may be prepared ahead of time by subject matter experts or stored on the spot by decision makers for later re-use.

ISDM utilizes templates in a variety of ways to support both problem definition and model elaboration:

- *Template library.* Users can manually browse or search the *Templates* library at any time. Templates can be any size, ranging from a single factor (i.e., node) to a network of conditions, decisions, and consequences, including default quantitative parameters (e.g., prior probabilities, conditional probabilities or impact weights, and importance weights on sources of value). Users can drag and drop any number of templates from the *Templates* library, and may then modify, specify, delete, or add individual factors and causal links, including links among templates and user created factors. Newly created templates may be stored in folders created by the user. Templates also contain key words or free text, which are used for matching and search.

- *Automatic recommendations based on total problem state*. ISDM matches stored templates against the attributes of an evolving decision problem, including initial information about the mission (type, location, relevant populations, etc.), plus (if model building is already underway) the current model text and semantic profile. Templates are retrieved and prioritized in order of their resemblance to the problem as it exists at any given time. Templates selected by the user appear in the model. In this example, the template shown in Figure 11 is recommended to the decision maker, based on the match between its rightmost node and the leftmost nodes in Figure 10.
- *Factor-specific recommendations on request*. For any factor in the existing model, right-click menus enable the user to view a list of factors that have previously been used as causes of the factor in any template, and similarly, a list of factors that have been used as effects of the factor in any template. The user can also view a list of multi-factor templates that include the factor. Factors or templates selected by the user from these lists appear in the model, with appropriate quantitative parameters and causal relationship to the original model factor. This feature prompts attention to causes of outcomes and consequences of actions.

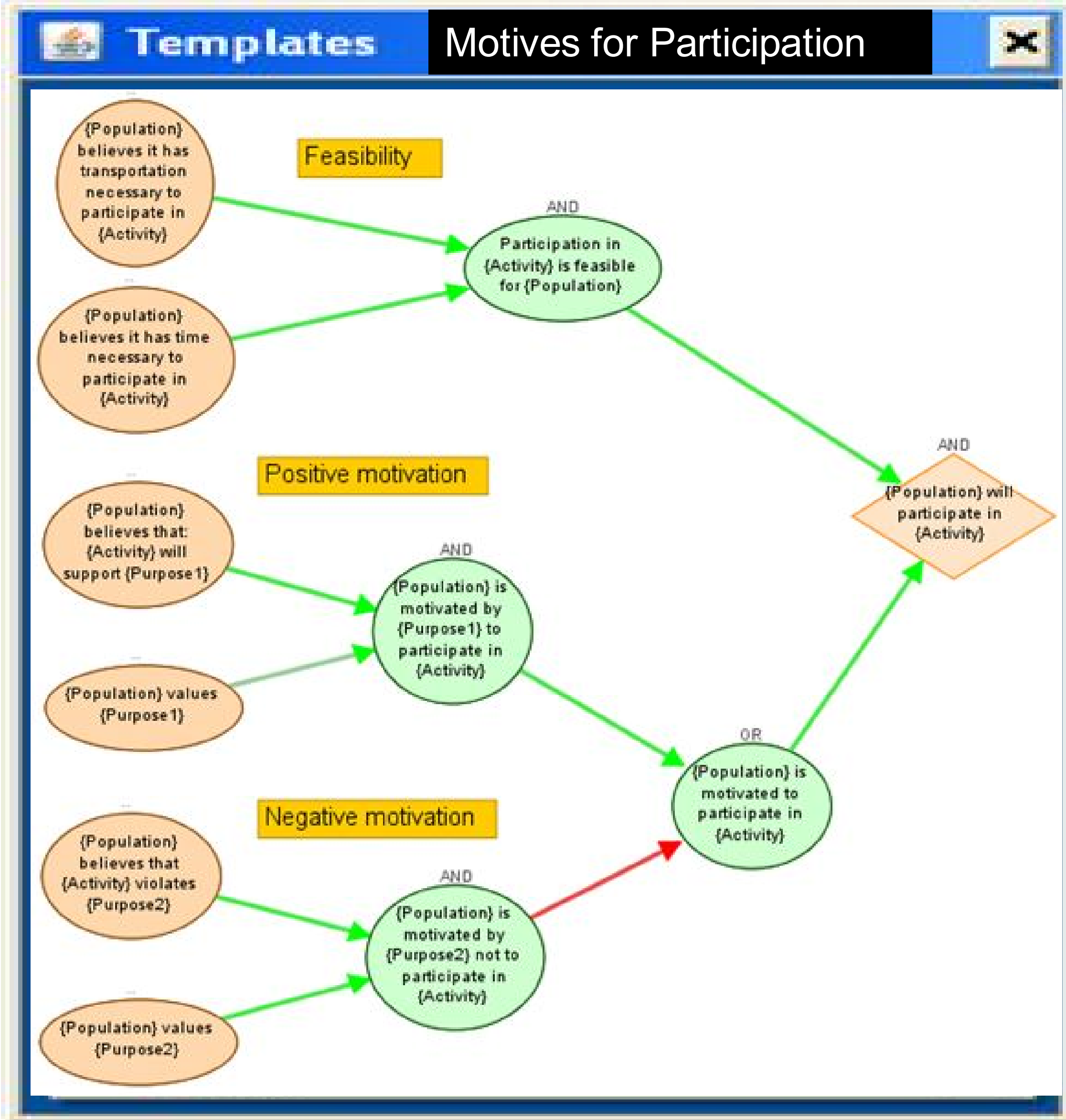

Figure 11. General template for forecasting participation in an activity by a population.

Additional features of ISDM templates significantly enhance power and flexibility: For example, as mentioned earlier, factors in templates may have *slots* marked by curly brackets enclosing an indexed category name, e.g., {Population} or {Purpose1} in Figure 11. When templates are added to the Model Composer workspace, users can fill in the slots with specific instances or lists of instances. Factors with slots represent predicates with open places for object reference,. Numerical category indices represent object identity constraints. All slots in a multi-factor template or decision model with the same category name plus numerical index are interpreted as referring to the same entity; e.g., {Purpose1} must refer to the same concept in all three of the factors where it appears in Figure 11. Slots with the same category name and different numerical indices (e.g., {Purpose1}, {Purpose2}) must refer to different entities. Categorized slots allow multi-factor templates to represent general knowledge, i.e., causal relationships among events involving object types. General templates are stored with quantitative parameters corresponding to prior and conditional probabilities for the categories in question; more accurate probabilities may be acquired over time for specific sub-categories or instances.

Figure 11 is a general template for degree of participation in an activity by members of a population, which has been automatically retrieved by matching to the initial problem formulation created by the user in Figure 9. The template in Figure 11 reflects a *theory of intentional actio*n, roughly corresponding to concepts and relationships that people tend to use when they construct explanations of their own or others' behavior (Malle, 2006; Baker, Saxe, & Tenenbaum, 2009; Cohen et al., 1996). It has factors representing goals, beliefs about means for achieving the goals, and beliefs about feasibility or opportunity with regard to means. Abstract templates of this sort can be used as starting points for detailed models in specific problems, or for more substantive templates.

The user can fill in the slots in Figure 11 to customize it to the current problem by selecting one or more items from drop down lists based on Catalog panels (see Figure 10), which contain elements from the category associated with the slot. The user can also type in new object names, which are automatically added to the Catalog under the relevant category. After slot fillers are specified, each original factor is replaced by an instance for each of the specified objects, preserving relevant causal connections and object identity constraints. The Catalog of types and instances in the relevant domain is modifiable; items can be added or deleted at any time.

Figure 12 illustrates this top down approach to model specialization and elaboration. It results directly from specifying a handful slots in Figure 11, without any other actions. The user has typed "Civic council" to fill the {Activity}slot, and "Safety" to fill the {Purpose2} slot, as a reason not to participate in the civic council. The user has filled in the {Purpose1} slot with a *list* of positive reasons, by browsing in the Catalog for the *Cultural Values* subcategory under *Purposes*, and selecting from a menu that includes family benefit, material benefit, social status, affiliation, obedience, tradition, achievement, personal development, moral principle, and so on. Because the user has checked off multiple Cultural values to fill the {Purpose1} slot, each factor containing a {Purpose1} slot is replaced by n factors, where n is the number of specified slot fillers. If more than one slot had been multiply specified, factors are created for the Cartesian product of the sets of specified slot fillers – subject to category and identity constraints between slots. Causal links among the original factors are replicated among the factors that result from this process, conditional on having the same slot fillers in slots that have the same category and object identity constraints. Figure 12 can now be saved to provide a more detailed alternative to Figure 11 in future problems that resemble the civic council decision.

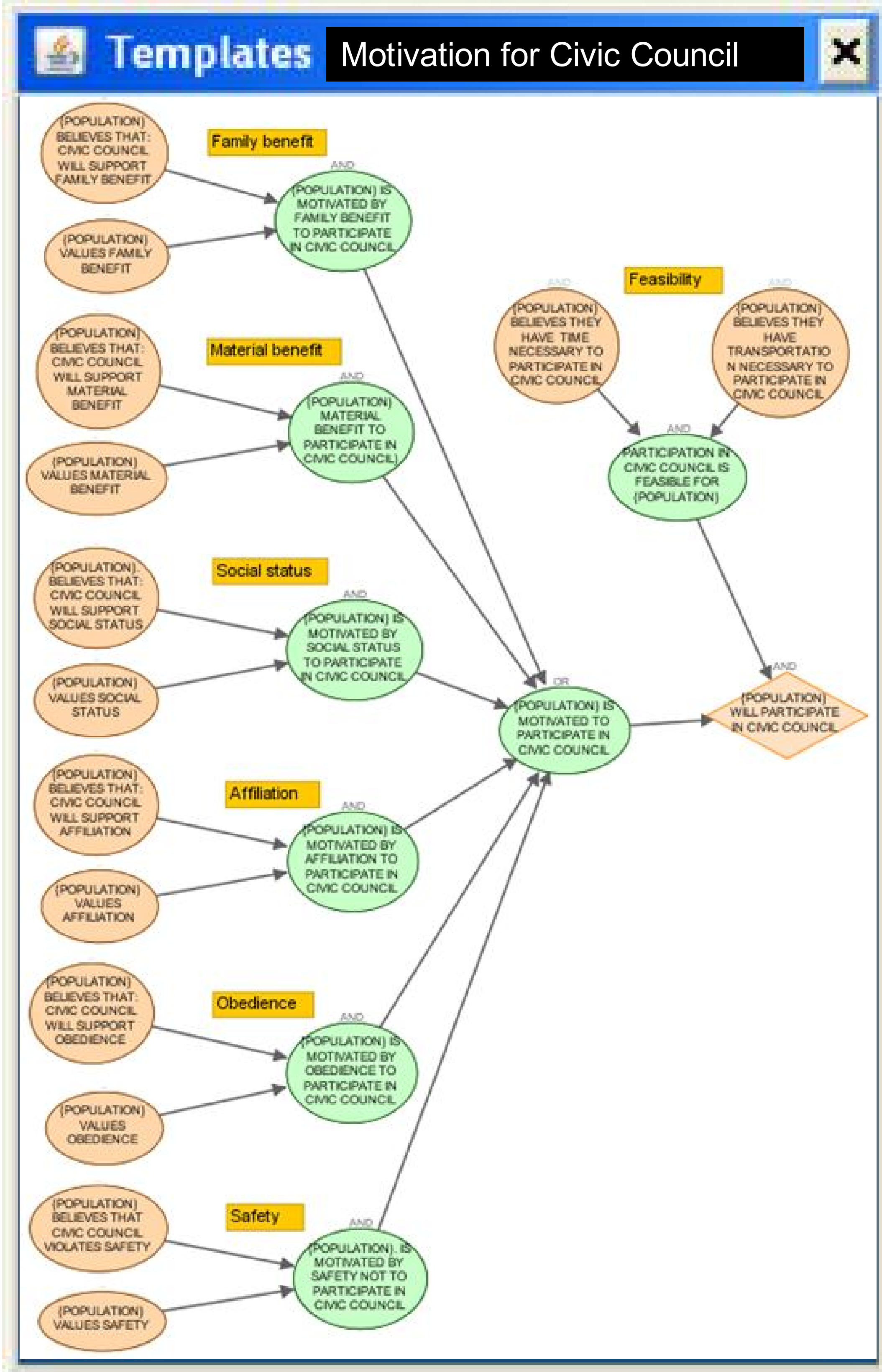

Figure 12. More detailed template created on the basis of general template in Figure 11.

In Figure 13, the user has dragged and dropped the more detailed template (Figure 12) into the model workspace and filled the {Population} slot by selecting multiple names from a list of populations, as in Figure 10A. Since every node in the template has this slot category, a separate model is generated for each specified population and placed in a distinct tab. Although they can be computed and viewed separately, the four models also function as part of a single comprehensive model – because each cultural model like the one in Figure 13 shares an element with the top-level decision model in Figure 10B. In the case of Figure 13, the shared element is “Sunni population will participate in Civic council.”

Another useful property of templates that facilitates the construction of larger models is that they are automatically *self-aggregating*. If two or more templates in the model make the same statement about the same objects, the factor only appears once in the model, inheriting the parents, children, and quantitative parameters from each template after testing for consistency. Note that the distinct cultural models (e.g., Figure 13) each share a node – e.g., Sunni pop. will participate in Civic council – with the top level decision model (Figure 10B). All instances of the shared node are treated as a single node in the underlying model comprising all the tabs, and provides the glue that binds the separate submodels to one another.

Self-aggregation can also support a more bottom-up approach to model or template specialization and elaboration. For example, Figure 14 begins with small, separate models for specific kinds of motivation for volunteering, which have research support: e.g., material benefits (upper left) and social status (upper right). Dragging them both into the same Model composer workspace leads to a connected network, unified on the common objective, “{Population} will participate in {Volunteering},” with a Noisy-Or operator. A detailed model of motivations for volunteering like Figure 12 could be built in this way from the bottom up and saved as a specialized elaboration of the more general template in Figure 11.

Aggregation is supported in a variety of ways. If two templates use different canonical relationships for the same factor, both sets of quantitative assessments are saved; the user can choose between them or add an intermediate variable to capture both relationships. The user can also manually merge factors by dragging and dropping one on top of the other and then editing the result. Smaller templates can be combined quickly to create larger patterns, or, conversely, larger patterns can be pared down to the elements relevant in a particular context.

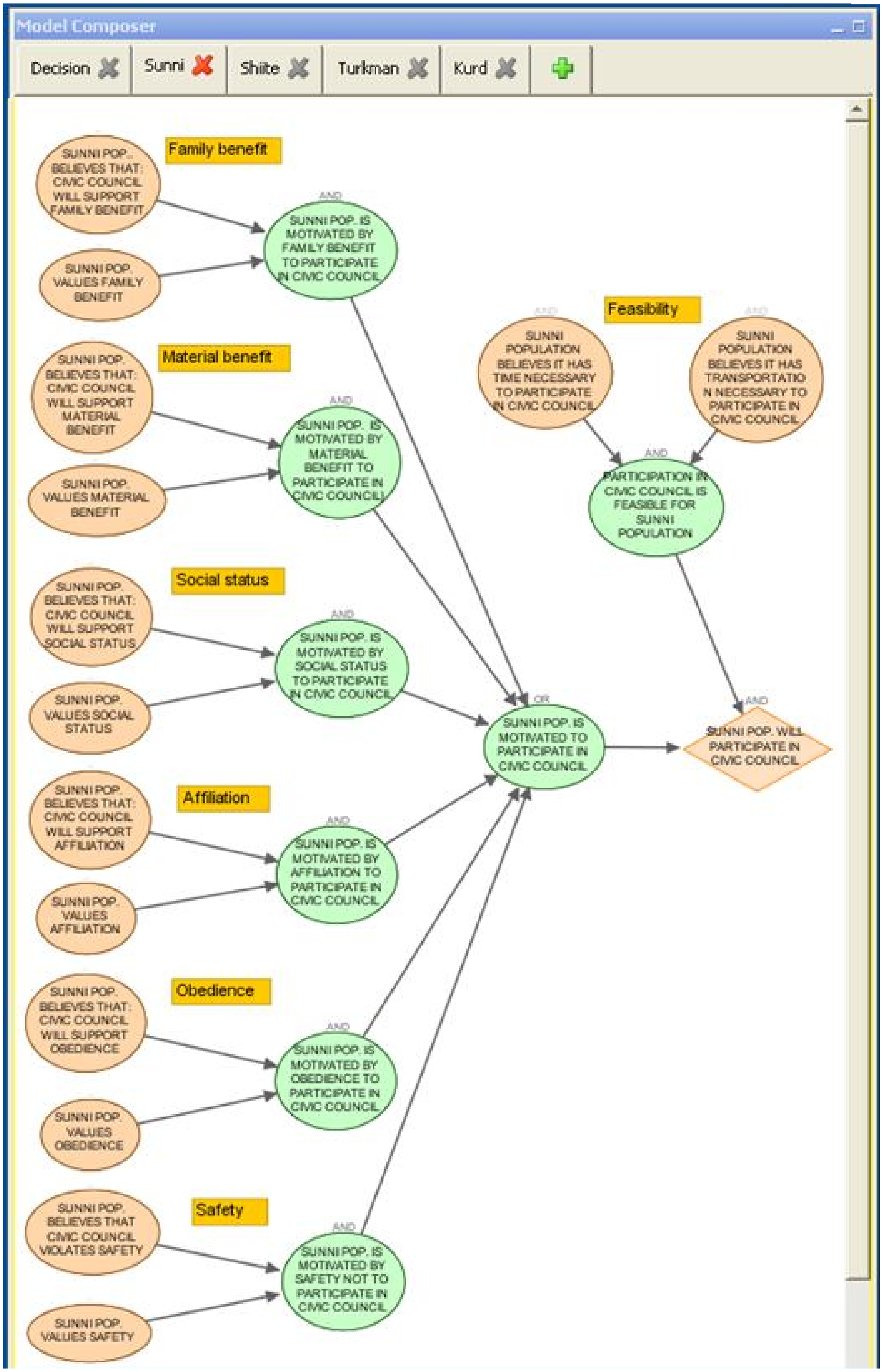

Figure 13. Illustrative result of filling slots for template in Figure 11.

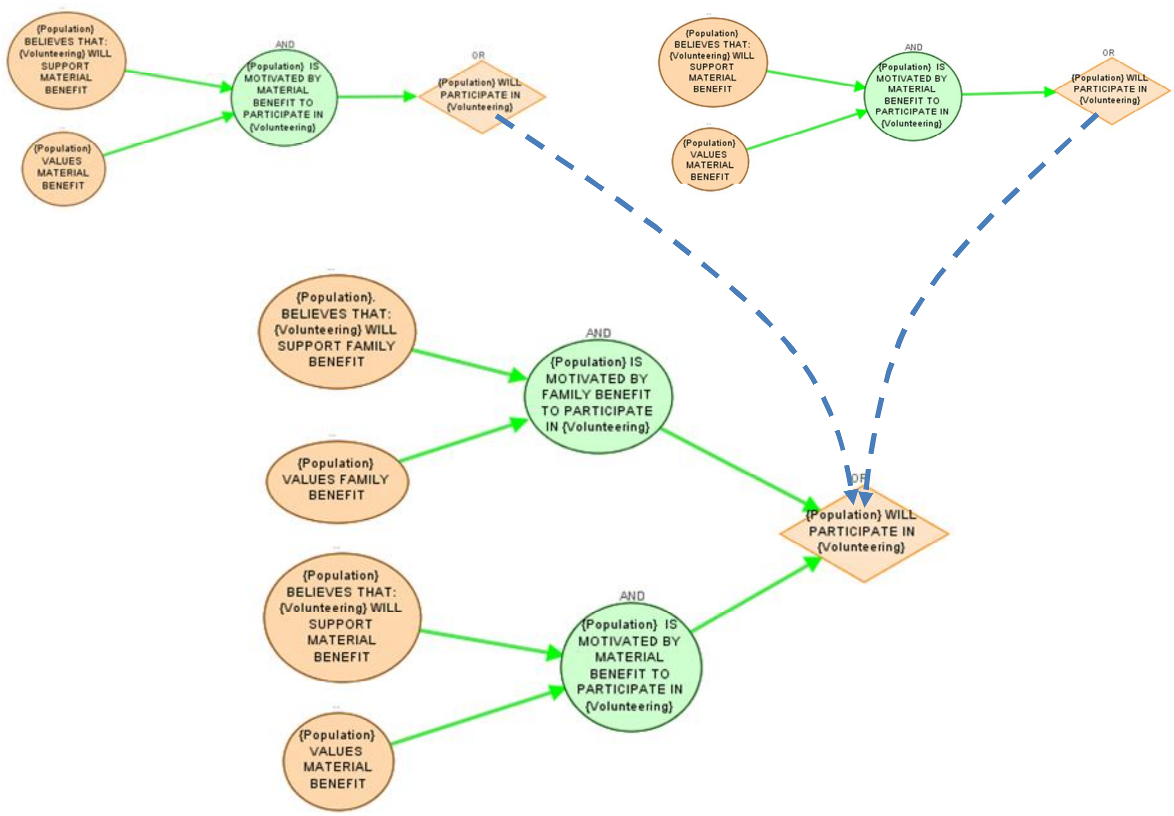

Figure 14. Merging identical factors from different templates.

These features in combination allow ISDM users to rapidly construct new models from prepared elements, even in novel circumstances. In relatively routine decisions, templates support decision making in the absence of substantial decision maker inputs. In less familiar situations, templates support users by kick-starting information selection and organization at the beginning of planning and by providing flexible building blocks for customized decision models. Users need only identify *new* factors, relationships or objects and adjust quantitative parameters that reflect *unusual* aspects of a problem.

A second function of templates is to store and accumulate substantive domain knowledge. Templates may have been developed ahead of time by subject matter experts, or they may have been saved by decision makers in past problems. Users can save new or modified models or model parts as templates for future use, expanding the template repertoire over time. Templates and models such as those in Figure 12 can be used to frame survey questions designed to pinpoint potential motivations for civic participation, and then to generate actions that encourage participation in accordance with basic social science research.

Figure 15 is another example of a detailed template for use in a different type of cross-cultural problem. It explains conflict among ethnic groups by means of alternative scenarios centered on different emotion terms, e.g., fear, anger, greed, resentment, and hatred, based on research by Collier & Hoeffler (2000), Lakoff & Kovecses (1987), and Peterson (2002).

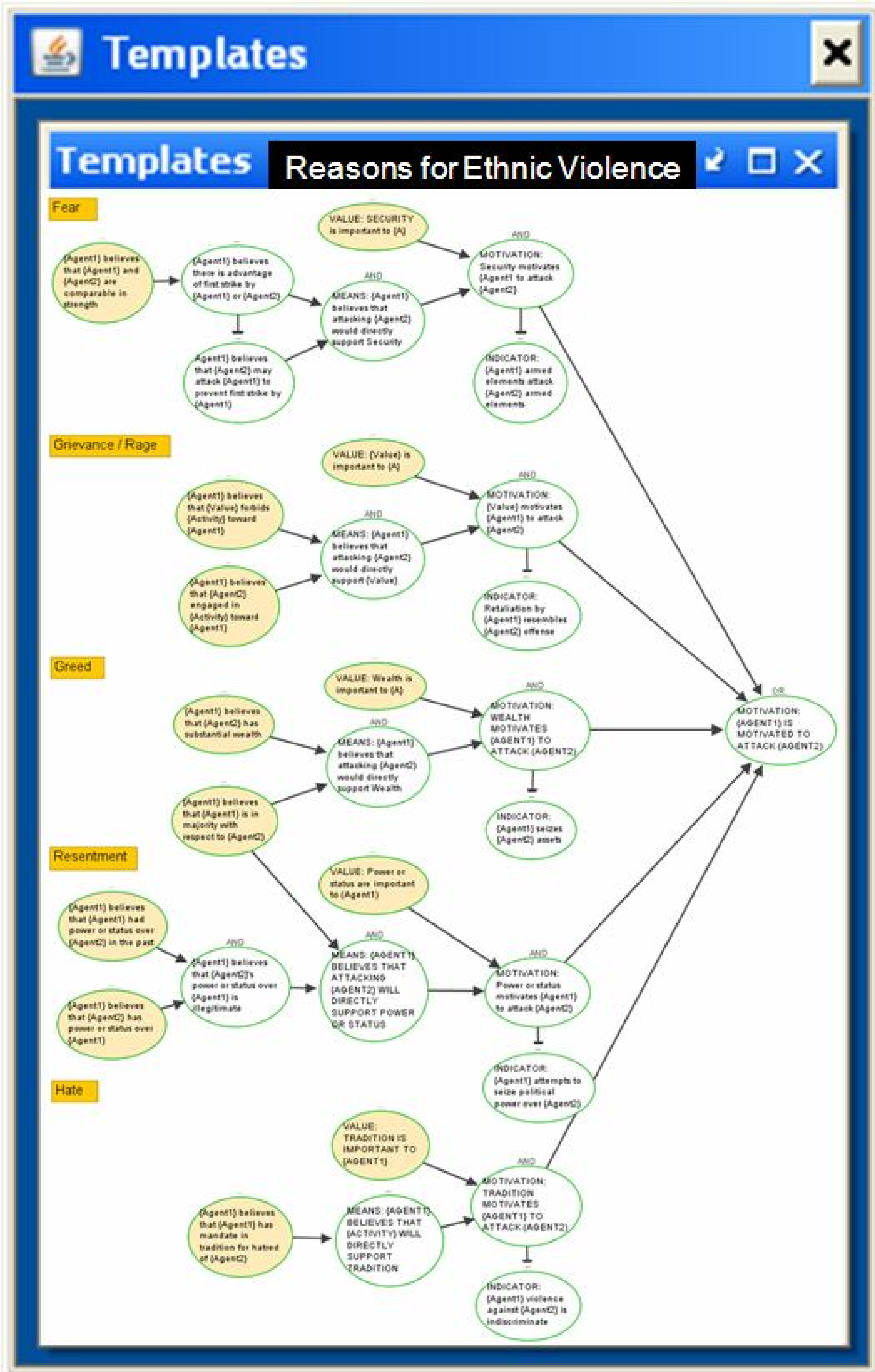

Figure 15. Illustrative template for explaining violence among ethnic groups.

Identify Information Requirements and Develop Questions

An ISDM dialog, the *Survey Composition Advisor*, supports two closely related aspects of questionnaire composition:

- Find the key uncertainties that will influence the decision by affecting its outcomes. As a decision model expands, it identifies *causal influences* on, and effects of, the critical situation factors. Each of these influences, if it is uncertain, is an additional information opportunity. ISDM generates a prioritized list of situation factors and parameters, which jointly determine outcomes, in order of value of information, i.e., the expected impact on utility of learning the actual state of the factor or value of the parameter. Such a prioritization is more reliable on subsequent iterations of modeling the same population. Early in the process, it will be based primarily on structural distance of a given indicator from utility.
- Retrieve or develop reliable survey questions regarding the state of the key causal influences. Templates associate situation factors and parameters, which jointly determine outcomes, with *indicators*, i.e., observable evidence regarding the state of the factor in question. When factors refer to human intentions, values, or beliefs, indicators may include differences in *behaviors* typically evoked by different states of the factor in question. For example, the template for ethnic violence shown in Figure 15 specifies behavioral indicators that discriminate among different motivations for attacking other groups; e.g., whether the attacks are indiscriminate or targeted, and if the latter, against military, political, or commercial elements of the other group. Indicators for these factors include counts of observed or published incidents of the behavior in question, analysis of published historical narratives, newspapers, internet commentary sites, direct judgment by area experts, and others. Most importantly for our purposes, indicators can also include answers to survey questions by the individuals whose beliefs, values, or intentions are of interest.

The Survey Composition Advisor offers advice and options to users and registers user responses by a process that applies whatever the current state of the model in the workspace. The starting point may be a single node, or a network like the ones in Figure 12 and Figure 15. The dialog follows a dual strategy: expanding the model until factors are arrived at for which indicators can be readily formulated, and formulating reliable indicators for the factors that are so identified:

1. Display situation factors in order of value of information. For each situation factor in the current model, check for the existence of prestored indicators. If found, highlight the relevant factors and provide access to a display of the indicators. Users can check off indicators that they wish to include in the survey.
2. If none are found, search paths from the situation factor forward to its effects and backwards to its causes and to other possible effects of those causes. Use both the model and templates that have not been added to the model (the latter is necessary for causal precursors of root nodes). Continue until prestored indicators are found or the path terminates. If factors with indicators are found among the causal precursors or effects, display the causal path or paths, highlight the factors associated with indicators, and provide access to a display of the indicators. If the user accepts the indicators on that path, and the path is not in the model, expand the model to include the factors and links in the path.

3. For factors without a satisfactory set of indicators, users can initiate a dialog regarding the best measurement method for that factor:
    a. Can the state of the factor be reliably judged by the respondent, i.e., the factor is relatively unambiguous in meaning, and the respondent has access to the relevant information, is likely to interpret it accurately, and is likely to report it in an unbiased manner? This category includes relatively straightforward demographic or other factual assertions and some reports of the respondent's own attitudes and beliefs. If yes, compose a question that asks the respondent directly for the state of the factor (yes/no, multiple choice), or degree of confidence or importance (numerical rating). Associate the question with the factor as an indicator. If the user is satisfied, mark the factor as *done*, and go to the next factor.
    b. Are there observable behaviors reliably correlated with the state of the factor? If yes, compose a question that asks about the frequency of the relevant behaviors from reliable observers. Associate the question with the factor as an indicator. If the user is satisfied, mark the factor as *done*, and go to the next factor.
    c. Does the factor refer at an abstract level that is unlikely to be interpreted unambiguously by respondents? If yes, identify more concrete questions that can be reliably judged by respondents and which, in aggregate, are likely to reliably predict the state of the factor. Associate each question with the factor as an indicator. (No single indicator need be conclusive.) If the user is satisfied, mark the factor as *done*, and go to the next factor.
    d. Is the state of the factor definable as a pattern of states of *other* factors across respondents? This category includes population traits like *Conformity*, *Collective Efficacy,* or *Mutual Trust*, all of which imply a concordance of responses within a socially inter-connected group of individuals. If yes, offer templates that capture patterns of responses across individuals in a population, and specify the relevant pattern of states of those factors as an Indicator of the original factor.

Thus, the model is elaborated until diagnostic questions about key factors have been found or formulated. Figure 16 illustrates the result of applying this process to the template in Figure 12.

The template in Figure 12 expresses each type of motivation to perform an activity by factors concerning (1) the *importance of values* and (2) corresponding beliefs that *the activity is a means of supporting the values*. The template also includes enabling conditions related to *feasibility*. An expansion of this kind is successful if it culminates in precursor factors for which survey questions can be formulated, which (a) produce reliable answers by directly querying states of the factor, or (b) produce reliable answers to a set of indirect questions whose answers are reliably correlated with the state of the factor.

For example, there are several ways to approach measurement of the *importance of values* associated with motives, which the Survey Composition Advisor will support:

- Ask survey respondent to directly rate the importance of each value (e.g., *security*, *achievement*, *family benefit*, *principle*, etc.) relative to the others in their life. A qualitative scale (e.g., high, medium, low) may be used.
- Formulate questions about more concrete preferences or feelings associated with the value, for which respondents can provide more reliable answers. For example, instead

of rating values, respondents would give yes-no answers to questions like: *I worry about my children's safety when they are not in the house*; *I prefer to travel in groups than alone in certain areas*, etc. Again, results can be factor analyzed. This is similar to the method used in personality inventories such as Meyers-Briggs.

- Associate each value with subsidiary values that are conceptually associated with it. Ask respondents to directly rate the relative importance of each subsidiary value. For example, instead of or in addition to rating *security* directly, respondents would rate the importance of reducing crime, reducing insurgent attacks, ability of children to walk to school without fear, ability to use public spaces without fear, and so on. Exploratory or confirmatory factor analysis can be used to identify or confirm the influence of the more general values. This approach is utilized in research on cultural values by cross-cultural psychologists (e.g., Hofstede, 2001: Schwartz, 1992).

The *final* step is to formulate a relationship between responses to survey questions and quantitative model parameters. The system provides default relationships for each type of response: When responses are ratings, the default is to convert qualitative ratings to numerical values and average across responses. When responses are agreement, the default is to use the proportion of *agree* responses.

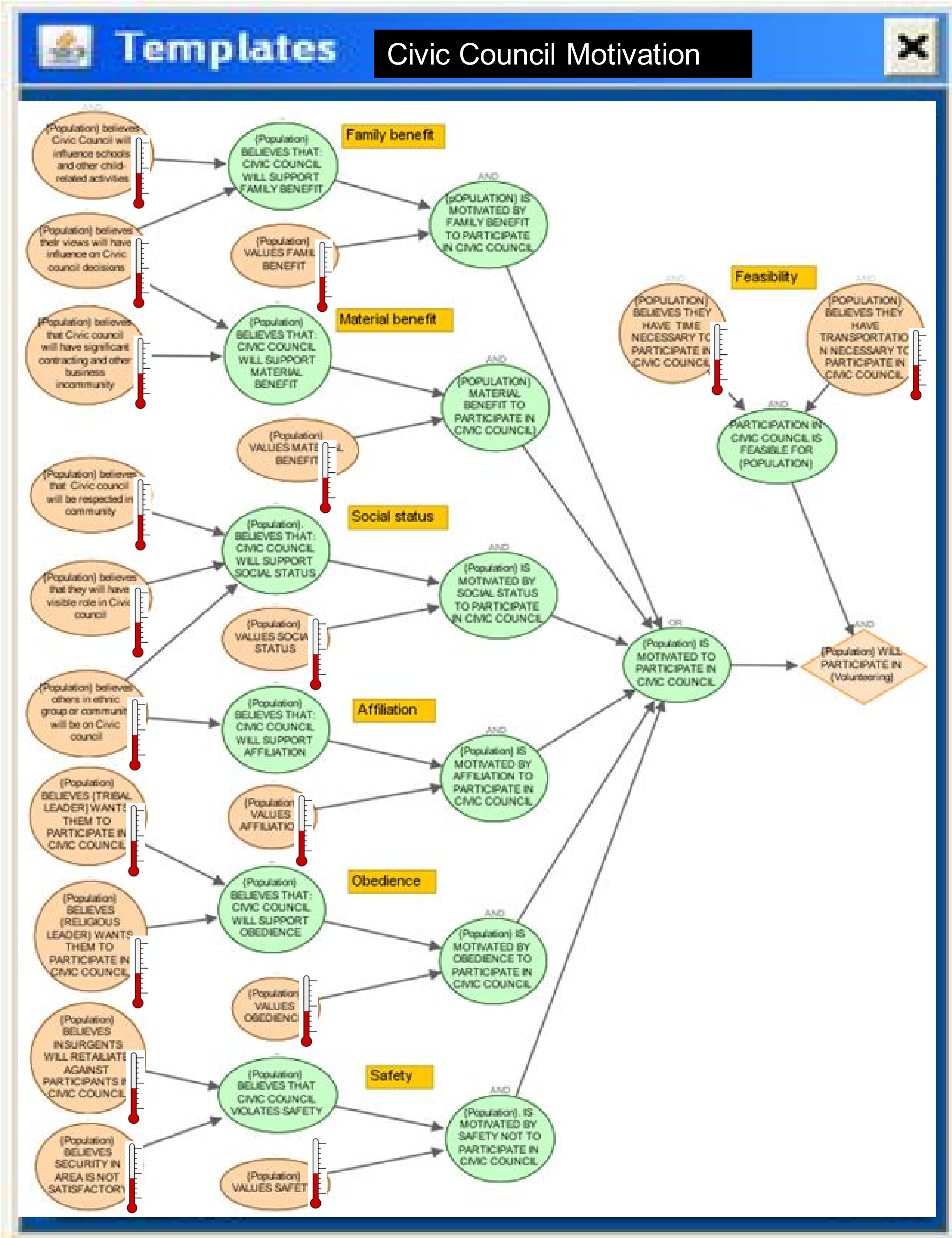

Figure 16. Template includes measurable factors and survey questions. Thermometers represent factors for which survey questions have been developed

Define Distinct Cultural models

If the user has predefined populations of interest (as in Figure 10), cultural consensus analysis may be used to identify subpopulations within each of the main groups. Alternatively, the analysis may be run on the total data set without presuppositions regarding significant divisions. Correlation with ethnic identify can be tested after emergent populations are identified by CCA based on survey responses.

In either case, ISDM displays a cultural model for each subpopulation identified by Cultural Consensus Analysis. In the civic council case, for example, different subpopulations might be susceptible to very different types of motivation, and there is no guarantee that such differences correlate strongly with independently verifiable attributes such as ethnicity, age, gender, socio-economic status. The following figures illustrate three subpopulations that are motivated quite differently in regard to participation in the civic council. Population 1, as shown by its cultural model in Figure 17, is inclined to join civic activities based on obedience to authority and affiliation with others in the same ethnic group or community. By contrast, the prospect of material benefit and increase in social status are the main drivers for Population 2, as shown by its cultural model in Figure 18. The cultural model for Population 3 (Figure 19) shows that their primary motivation is family benefit. All three populations, however, are negatively motivated by lack of safety. Each model includes the parts of the civic participation template most relevant to the population, plus the results of expansion to identify measurable factors.

The contrasting cultural models suggest different ways to encourage participation in each of the three populations – independent of predefined distinctions (e.g., Sunni, Shia, Kurds, and Turkmen). Possible interventions for each subpopulation are also displayed in the cultural models. These may be created by the user and then saved as part of the complete template, as illustrated in Figure 20, for automatic retrieval in future cases.

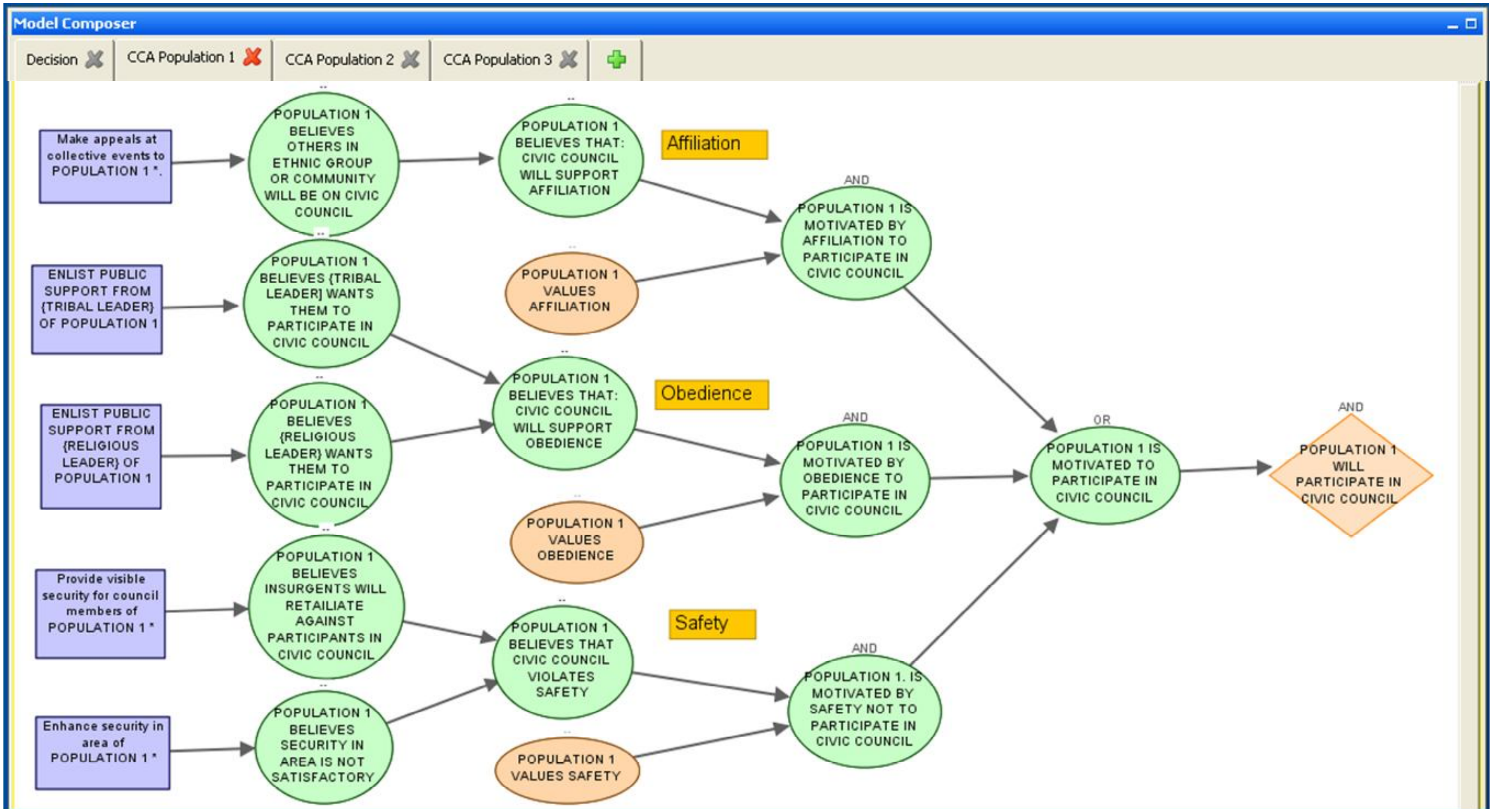

Figure 17. Cultural model for population 1 (illustrative). Rectangular nodes are possible interventions.

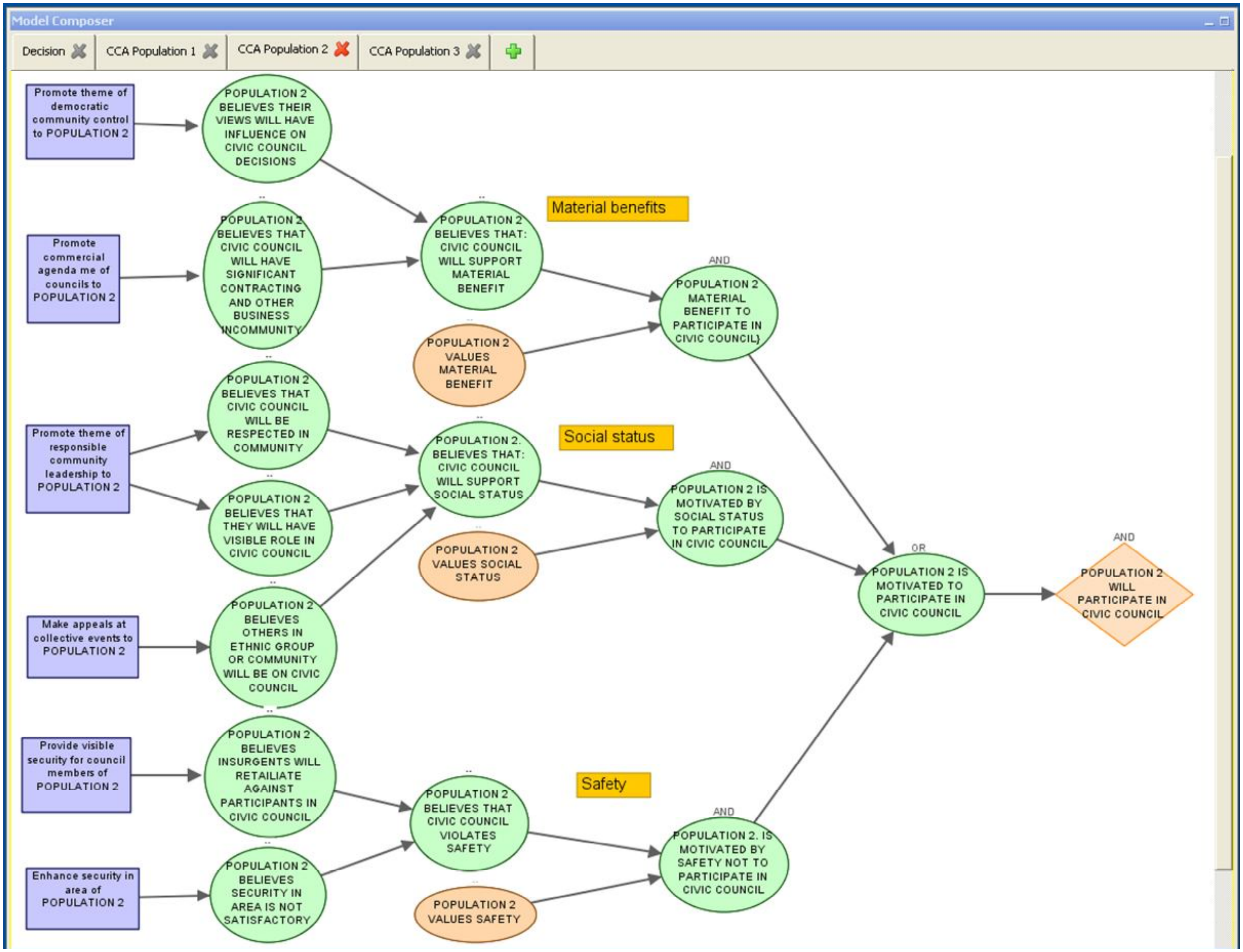

Figure 18. Cultural model for Population 2 (illustrative).

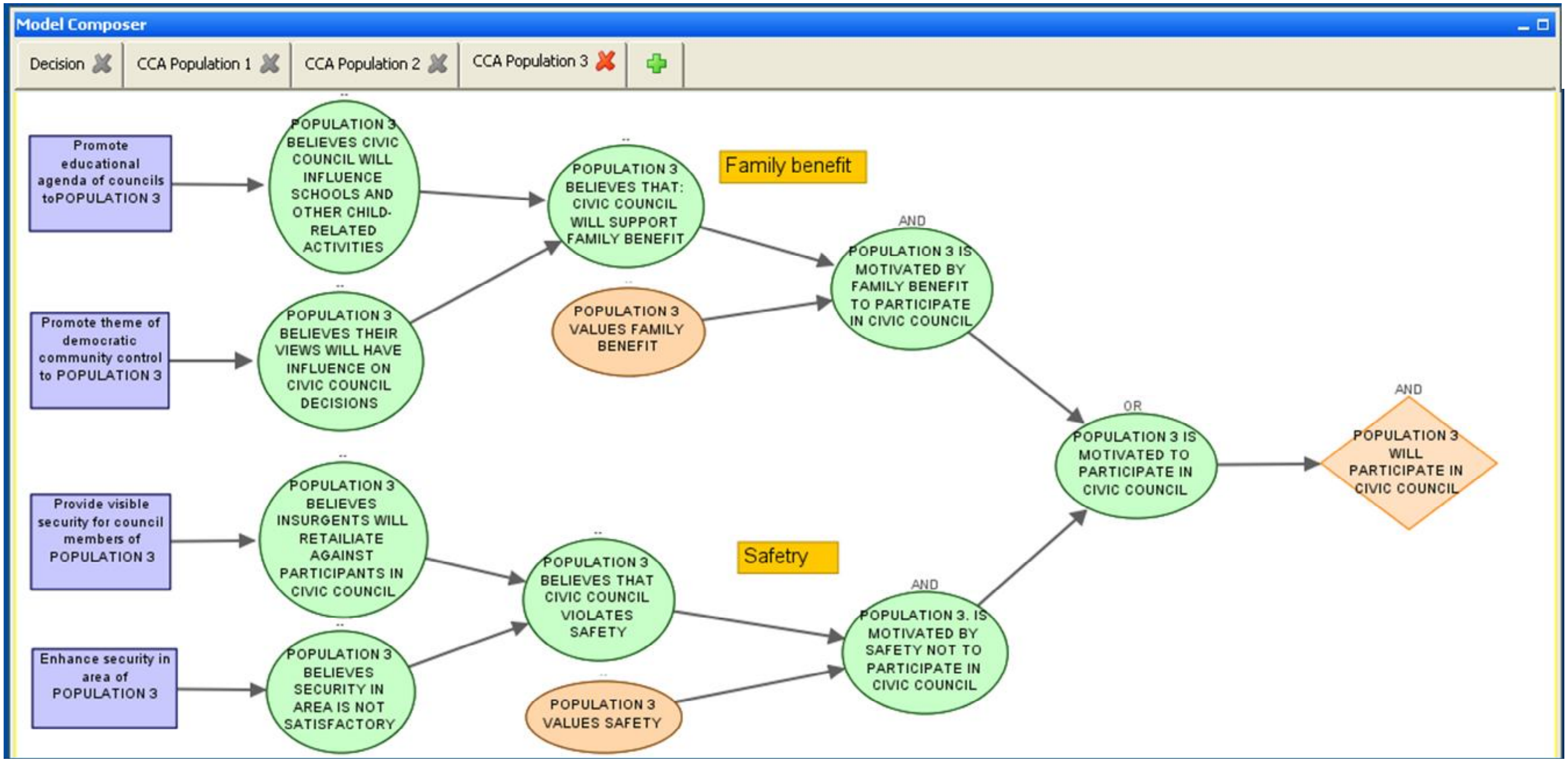

Figure 19. Cultural model for Population 3 (illustrative).

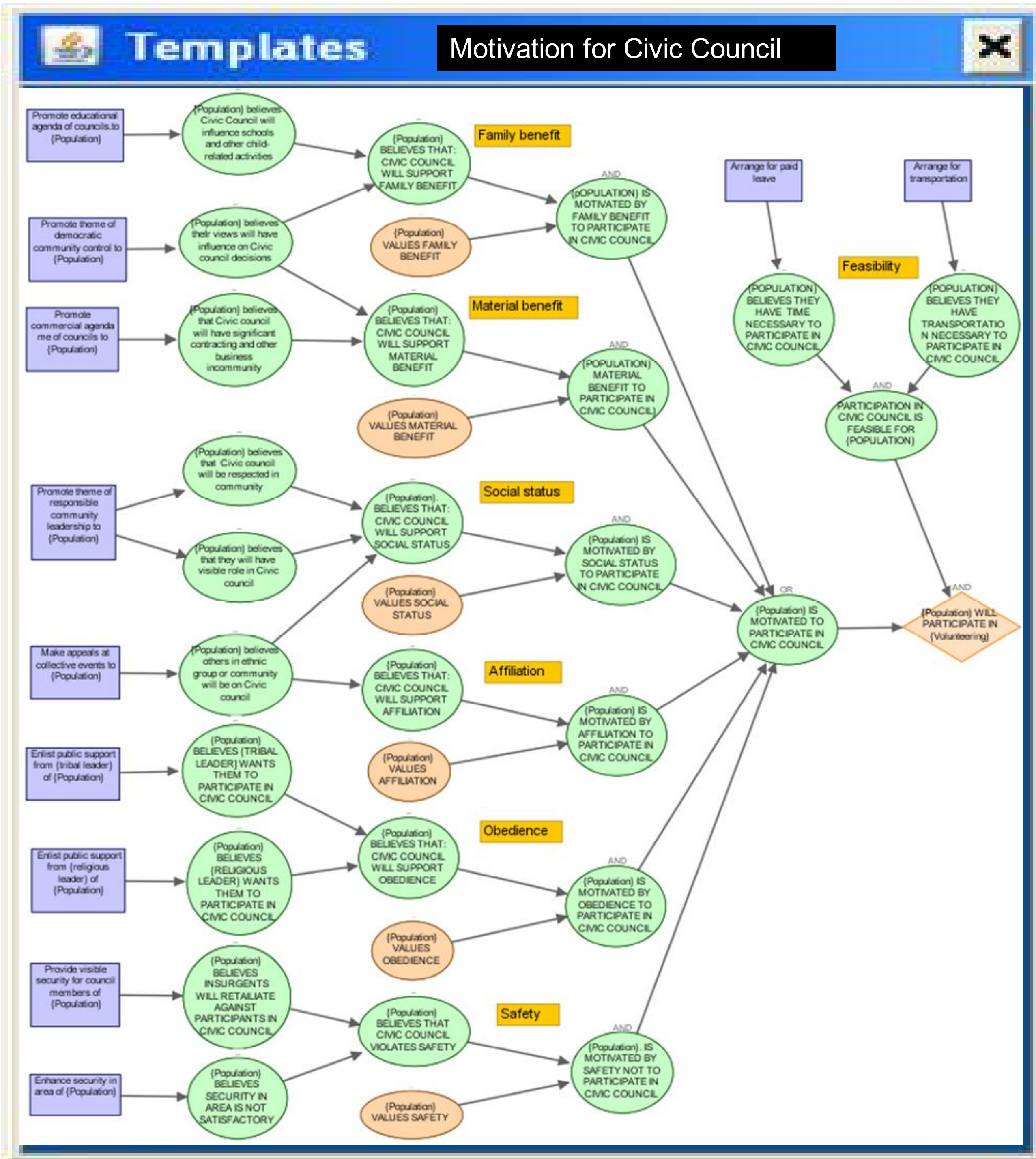

Figure 20. A complete templates for the civic council problem, which includes possible interventions that influence the states of key factors.

Integrate Cultural Models into Decision Model

. ISDM can incorporate cultural models as components of normative decision models, to forecast culturally moderated outcomes of possible courses of actions. Cultural models for each CCA-identified population are displayed in distinct tabs, as illustrated in the previous section. In addition, ISDM expands the initial decision model shown in Figure 9 by inserting the CCA-identified populations in the open {Population} slot. The distinct cultural models and the top-level decision model are automatically linked based on their shared factors, i.e., participation by a particular population in the civic council. The top level decision model display also includes the interventions identified as appropriate possibilities for each distinct population. When an operation applies to more than one population, users decide whether it can be applied to them separately; if not, the operation appears only once in the top-level decision model. The top level model also enables users to estimate costs on one or more dimensions for the primary decision (whether or not to encourage the formation of civic councils) and interventions designed to facilitate its success. The result is illustrated in Figure 21.

ISDM uses the integrated model to identify the highest scoring course of action, i.e., combination of actions including the original decision and specific interventions, taking both costs and benefits into account. Figure 22 displays the three best combinations of actions in three columns, from left to right. In this example, the best combination includes all the actions except "Enlist public support from {Tribal leader} of Population 1" – for which expected costs exceed benefits. The second best course of action includes this intervention.

Figure 22 displays the actions in rows, prioritized from top to bottom in terms of marginal contribution to the COA in the course of action in the column selected by the user. The user can add and subtract actions from any course of action by checking or unchecking the cell on the row for the action in the column for the course of action.

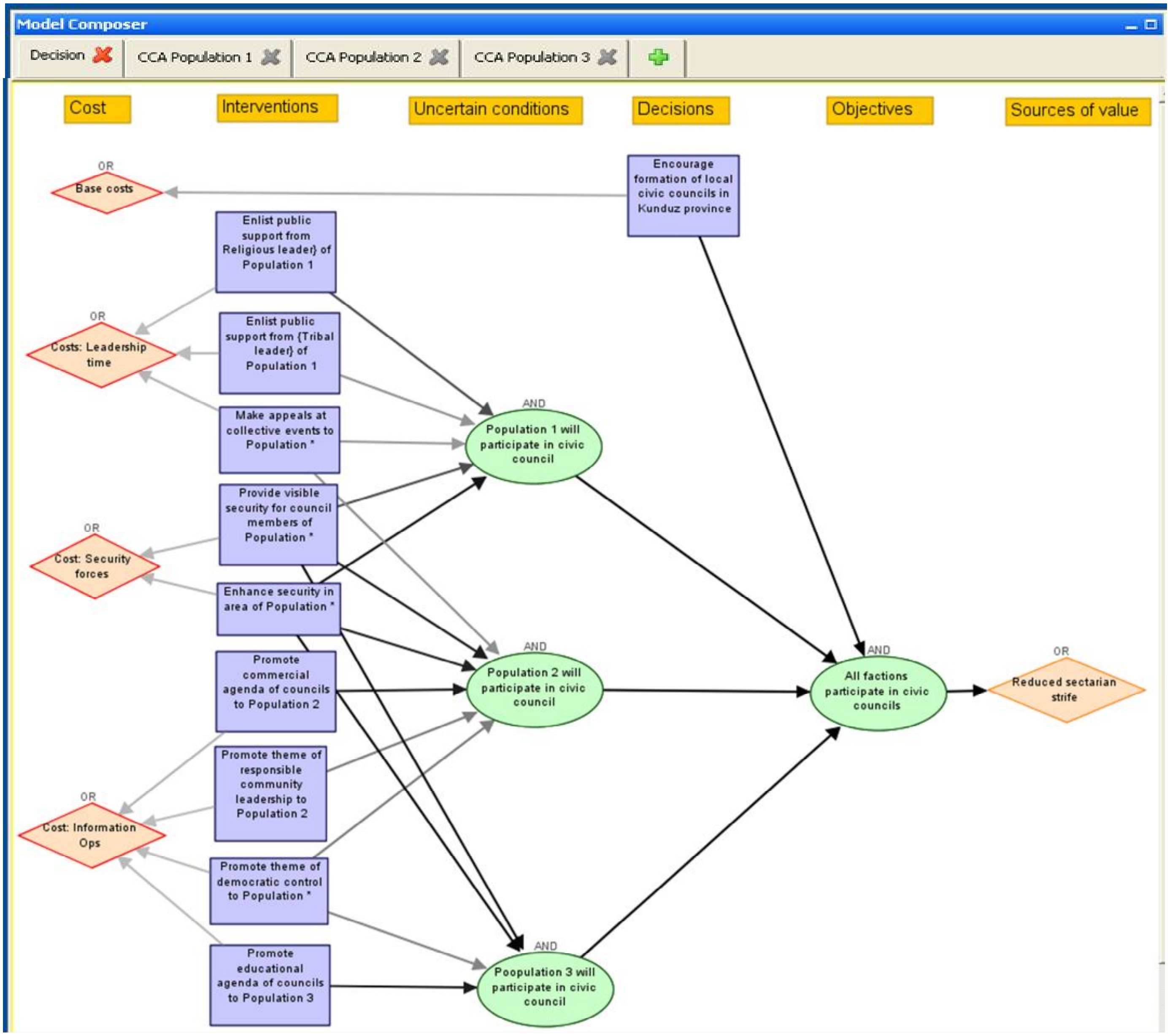

Figure 21. Decision model view showing interventions associated with different populations and their costs.

Figure 22. Ranking of candidate tasks and courses of action.

CONCLUSION

Figure 23 provides an overview of how the three ISDM technologies into fit together, both in terms of their individual functions and their synergistic interaction.

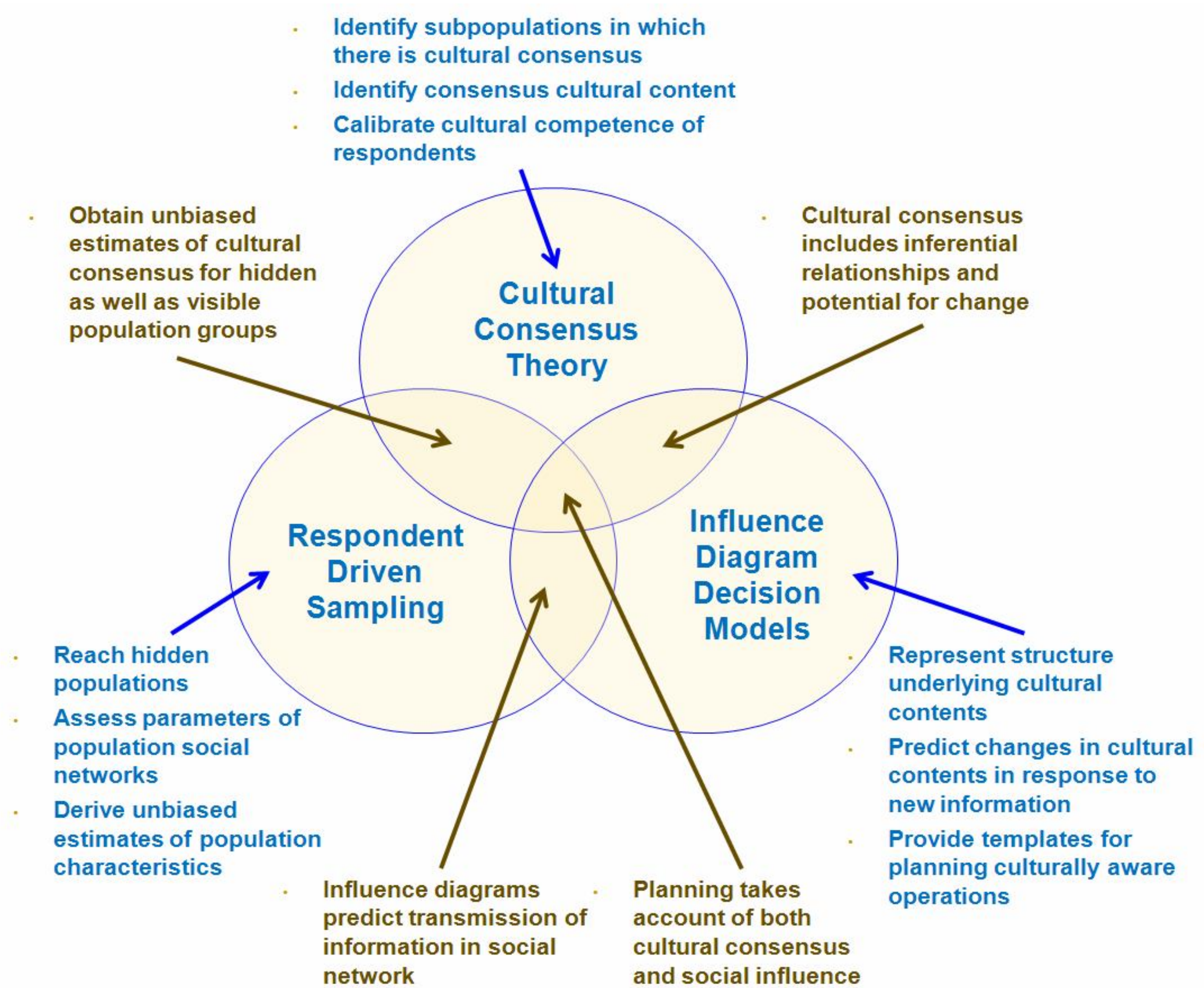

Figure 23. Relationships among the three technologies integrated in ISDM concept. Areas of overlap provide synergies.

ISDM is a software product that provides guidance and automated support in collecting data, analyzing it to accurately identify and estimate the size of distinct sub-populations, determine the crucial cultural elements that differentiate subpopulations from one another, and accurately differentiate respondents in terms of social network centrality and representativeness of their beliefs and values within each relevant subculture. It introduces an integration of RDS and CCA to efficiently collect the most useful data and draw valid statistical inferences about the beliefs, values, and behaviors of a population and its constituent sub-populations, even when its members are unknown in advance, hard to identify, or unwilling to be sampled, and if sampled, are prone to give conflicting views – while investing only a tiny fraction of the time and money required by conventional representative sampling.

ISDM utilized influence diagrams as an appropriate representation format for cultural knowledge, which is (i) general and flexible enough to avoid cultural bias or preconceptions, (ii) well-defined and precise enough to support RDS and CCA analysis, (iii) a perspicuous encapsulation of culturally specific reasoning processes, and (iv) logically valid and computationally efficient as a medium to support real-time simulation and decision making. ISDM makes cultural models available as decision templates that can be combined into models and used by planners to simulate outcomes and select a course of action.

REFERENCES

Atran, S., Medin, D. L., & Ross, N. O. (2005). The cultural mind: Environmental decision making and cultural modeling within and across populations. *Psychological Review*, *112*(4), 744-776.

Baker, C., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*, 329-340.

Batchelder, W. H., & Romney, A. K. (1988). Test theory without an answer key. *Psychometrika*, *53*(1), 71-92.

Benedict, R. (1946). *The chrysanthemum and the sword: Patterns of Japanese culture*. Boston: Houghton Mifflin.

Boas, F. (1963). *The mind of primitive man: Revised Edition*. New York: Collier Books.

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, *27*, 55-71.

Brown, R. V., & Paschoud, P. (2005). *Rational choice and judgment: Decision analysis for the decider*. Hoboken, NJ: Wiley.

Bussell, H., & Forbes, D. (2001). Understanding the volunteer market: The what, where, who, and why of volunteering. *International Journal of Nonprofit and Voluntary Sector Marketing*, *7*(3), 244-257.

Clemen, R. T., & Reilly, T. (2001). *Making Hard Decisions with Decision Tools*. Pacific Grove, CA: Duxbury.

Cohen, M. S., & Freeling, A. N. S. (1981). *The impact of information on decisions: command and control system evaluation*. Falls Church, VA: Decision Science Consortium, Inc.

Collier, P., & Hoeffler, A. (2000). *Greed and Grievance in Civil War*: The World Bank Development Research Group.

D'Andrade, R. (1995). *The Development of Cognitive Anthropology*. Cambridge, UK: Cambridge University Press.

D'Andrade, R., & Strauss, C. (1992). *Human Motives and Cultural Models*. Cambridge, UK: Press Syndicate of the University of Cambridge.

de Finetti, B. (1964). Foresight: Its logical laws, its subjective sources. In J. Kyburg, H.E. & H. E. Smokler (Eds.), *Studies in Subjective Probability*. New York: Wiley.

Diamond, J. (1997). *Guns, Germs, and Steel: The Fates of Human Societies*. New York: Norton.

Diez, F., & Druzdzel, M. (2005). *Canonical probabilistic models for knowledge engineering*. Madrid: CISIAS, UNED.

Edwards, W., Miles, R. F., Jr., & von Winterfeldt, D. (Eds.). (2007). *Advances in Decision Analysis: From Foundations to Applications*. Cambridge: Cambridge University Press.

Ericsson, K. A., & Smith, J. (Eds.). (1991). *Toward a general theory of expertise: Prospects and limits*. Cambridge UK: Cambridge University Press.

Garro, L. C. (2000). Remembering what one knows and the construction of the past: A comparison of Cultural Consensus Analysis and cultural schema theory. *Ethos*, *28*(3), 275-319.

Geertz, C. (1973). *Interpretation of Cultures*. New York: Basic Books.

Glaser, R. (1989). Expertise and learning: how do we think about instructional processes now that we have discovered knowledge structures?. In &. K. K. D. Klahr (Ed.), *Complex Information Processing: The Impact of Herbert A. Simon*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Grannis, R. (2009), From the Ground Up: Translating Geography Into Community Through Neighbor Networks, Princeton University Press, Princeton, New Jersey

Grannis, R. (2010), "The Contagion of Cheating," *American Journal of Sociology*, In Press

Grant, K. L., & Miller, M. L. (2004). A cultural consensus analysis of marine ecological knowledge in the Solomon Islands. *SPC Traditional Marine Resource Management and Knowledge Information Bulletin*(17), 3.

Handwerker, W. P. (2008). The construct validity of cultures: Cultural diversity, culture theory, and a method for ethnography. *American Anthropologist*, *104*(1), 106-122.

Harris, M. (1999). *Theories of Culture in Postmodern Times*. Walnut Creek, CA: AltaMira Press.

Heckathorn, D. D. (1997). Respondent-Driven Sampling: A new approach to the study of hidden populations. *Social Problems*, *44*(2), 174-199.

Heckathorn, D. D. (2002). Respondent-Driven Sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems*, *49*(1), 11-34.

Heckathorn, D. D. (2007). Extensions of Respondent-Driven Sampling: Analyzing continuous variables and controlling for differential recruitment. *Sociological Methodology*, *37*, 151-208.

Heckerman, D., & Breese, J. S. (1996). Causal Independence for Probability Assessment and Inference Using Bayesian Networks. *IEEE Transactions on Systems, Man, and Cybernetics. Part A: Systems and Humans*, *26*(6), 826-831.

Henrion, M. (1989). Some practical issues in constructing belief networks. In L. Kanal & J. Lemmer (Eds.), *Uncertainty in Artificial Intelligence 3* (pp. 163-71). Amsterdam: Elsevier Science Publishers.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage Publications.

Holland, D., & Quinn, N. (Eds.). (1987). *Cultural Models in Language and Thought*. Cambridge UK: Cambridge University Press.

Howard, R. A., & Matheson, J. (1989). Influence Diagrams. In R. A. Howard & J. E. Matheson (Eds.), *The Principles and Applications of Decision Analysis* (Vol. II, pp. 719-62). Menlo Park, CA: Strategic Decisions Group.

Jeffrey, R. C. (1965/1983). *The logic of decision*. Chicago, IL: University of Chicago Press.

Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*. New York: Springer.

Jensen, F., Jensen, R. V., & Dittmer, S. L. (1994). From influence diagrams to junction trees. In R. de Mantaras & D. Poole (Eds.), *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence* (pp. 367-73). San Francisco: Morgan-Kaufman.

Karabatsos, G., & Batchelder, W. H. (2003). Markov chain estimation for test theory without an answer key. *Psychometrika*, *68*(3), 373-389.

Keeney, R. & Raiffa, H. 1976, *Decisions with multiple objectives: Preferences and value tradeoffs*, Wiley, New York.

Lakoff, G., & Kovecses, Z. (1987). The Cognitive Model of Anger Inherent in American English. In D. Holland & N. Quinn (Eds.), *Cultural Models in Language and Thought* (pp. 195-221). Cambridge UK: Cambridge University Press.

Larkin, J., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. *Science*, *20*(208), 10-335-1342.

Laskey, K., Cohen, M.S., & Martin, A.W. (1989). Representing and eliciting knowledge for uncertain evidence and its implications. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(3).

LaValle, I. H. (1968). On cash equivalents and information evaluation in decisions under uncertainty. *American Statistical Association Journal*, 63, 252-290.

LaValle, I. H. (1980). On value and strategic role of information in semi-normalized decisions. *Operations Research*, 28, 129-138.

Leddo, J., & Cohen, M.S. (1989). *Cognitive structure analysis: A technique for eliciting the content and structure of expert knowledge*. Proceedings of 1989 Al Systems in Government Conference, McLean, VA.

Lipshitz, R., & Cohen, M. S. (2005). Warrants for prescription: Analytically and empirically based approaches to improving decision making. *Human Factors*, *47*(1), 121-130.

Luce, R. D., & Raiffa, H. (1957/1989). *Games and Decisions: Introduction and critical Survey*: Dover.

Malinowski, B. (1984). *Magic, Science, and Religion and Other Essays*. Westport CT: Greenwood Press.

Malle, B. F. (2006). *How the Mind Explains Behavior: Folk Explanations, Meaning, and Social Interaction*. Cambridge, MA.: The M.I.T. Press.

Marshall, K. T., & Oliver, R. M. (1995). *Decision Making and Forecasting*. New York: McGraw-Hill.

Mead, M. (1928). *Coming of Age in Samoa*. New York: Morrow.

Mueller, S. T., & Veinott, E. S. (2008). Cultural Mixture Modeling: Identifying Cultural Consensus (and Disagreement) using Finite Mixture Modeling. *ARA Technology Review*, *4*, 39-45.

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently...and why*. New York: The Free Press.

Pearl, J. (1989). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo CA: Morgan Kaufmann Publishers

Pearl, J. (2000). *Causality: models, reasoning, and inference*. Cambridge UK: Cambridge University Press.

Peterson, R. D. (2002). *Understanding Ethnic Violence: Fear, hatred, and Resentment in Twentieth-Century Eastern Europe*. Cambridge: Cambridge University Press.

Radcliffe-Brown, A. (1965). *Structure and Function in Primitive Society*. New York: MacMillan.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty*. Reading, MA: Addison-Wesley Publishing.

Reyes-Garcia, V., Byron, E., Vadez, V., Godoy, R., Apaza, L., Limache, E. P., et al. (2003). Measuring culture as shared knowledge: Do data collection formats matter? Cultural knowledge of plant uses among Tsimane' Amerindians, Bolivia. *Field methods*, *15*(2), 1-22.

Romney, A. K., S. C. Weller, and W. H. Batchelder. (1986). "Culture as consensus: A theory of culture and accuracy." *American Anthropologist* 88:313-338.

Ross, N. (2004). *Culture and cognition: Implications for theory and method*. Thousand Oaks: Sage Publications.

Ross, N., & Medin, D. L. (2005). Ethnography and Experiments: Cultural Models and Expertise effects elicited with experimental research techniques. *Field methods*, *17*(2), 131-149.

Salganik, M. J., & Heckathorn, D. D. (2004). Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology*, *34*, 193-239.

Savage, L. J. (1954/1972). *The foundations of statistics*. New York, NY: Dover Publications.

Schoenfeld, A., & Herrman, D.J. (1982). Problem perception and knowledge structure in expert and novice mathematical problems solvers. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *8*, 484-494.

Schwartz, S. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in experimental social psychology* (Vol. 25, pp. 1-65). San Diego: Academic Press.

Shachter, R. D. (1986). Evaluating influence diagrams. *Operations Research*, *34*(6), 871-882.

Shachter, R. D. (2007). Model building with belief networks and influence diagrams. In W. Edwards, R. F. Miles, Jr. & D. von Winterfeldt (Eds.), *Advances in Decision Analysis:*

From Foundations to Applications (pp. 177-201-446). Cambridge: Cambridge University Press.

Shenoy, P. (1992). Valuation-based systems for Bayesian decision analysis. *Operations Research*, *40*, 463-484.

Shore, B. (1996). *Culture in Mind: Cognition, Culture, and the Problem of Meaning*. Oxford, UK: Oxford University Press.

Sieck, W. R., & Rasmussen, L. J. (2008). *Cultural network analysis: Characterizing target audience cognition*. Paper presented at the Proceedings of Phoenix Challenge, Phoenix, AZ.

Sloman, S. A. (2005). *Causal Models: How People Think about the World and its Alternatives*. Oxford: Oxford University.

Smith, J. B., & Batchelder, W. H. (2008). Assessing individual differences in categorical data. *Psychonomic Bulletin & Review*, *15*(4), 713-731.

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. Oxford, UK: Blackwell.

Strauss, C., & Quinn, N. (1997). *A Cognitive Theory of Cultural Meaning*. Cambridge, UK: Cambridge University Press.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. Cambridge, UK: Cambridge University Press.

Watson, S. R., & Buede, D. M. (Eds.). (1987). *Decision synthesis: The principles and practice of decision analysis*. Cambridge, UK: Cambridge University Press.

Watts, D. (1999). *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton NJ: Princeton University Press.

Weiser, M., Wolfe, C. R., Lawrence, J. A., & Engle, R. A. (1983). Programming problem representation in novice and expert programmers. *International Journal of Man-Machine Studies*, *19*, 391-398.

Weller, S. C. (2007). Cultural Consensus Analysis: Applications and frequently asked questions. *Field methods*, *19*(4), 339-368.

Weller, S. C., & Mann, N. C. (1997). Assessing rater performance without a "gold standard" using consensus theory. *Medical Decision Making*, *17*(1).

Weller, S. C., & Romney, A. K. (1988). *Systematic Data Collection*. Newbury Park, CA: Sage.

White, L. A. (2005). *The Science of Culture: A Study of Man and Civilization*: Percheron Press.